LIMBURGS
UNIVERSITAIR
CENTRUM
Faculteit Wetenschappen

Non-Linear Models for Multivariate Repeated Ordinal Data

P.J. Lindsey

Promotor
Prof. Dr. G. Molenberghs

April 27, 2001

Contents

Notation and Symbols

Vectors are bold lower case and matrices bold upper case. An exception is the response variable, y, where bold upper case means the vector random variable. Primes denote derivatives of a function and T the transpose of a vector or matrix.

$\mathbf{R}$	– the real line
$\mathbf{R}^m$	– the Euclidean space of dimension m
y	– observed response variable
$\mathbf{y}$	– vector of observed response variables
Y	– random variable
$\mathbf{Y}$	– vector of random variable
$G(y)$, s	– arbitrary function of y
$G_{\mathrm{M}}(\mathbf{y})$	– arbitrary multivariate function of $\mathbf{y}$
$f(y)$	– probability density function
$f(y\vert\cdot)$	– conditional density
$f_{\mathrm{m}}(y)$	– mixture density
$f_{\mathrm{m}}(y\vert\cdot)$	– conditional mixture density
$f_{\mathrm{B}}(y_1, y_2)$	– bivariate distribution
$f_{\mathrm{M}}(\mathbf{y})$	– multivariate distribution
F, $\widetilde{F}$, $F(y)$, $\widetilde{F}(y)$	– cumulative distribution function
$F(y\vert\cdot)$	– conditional cumulative distribution function
$F_{\mathrm{m}}(y)$	– mixture cumulative distribution function
$F_{\mathrm{B}}(y_1, y_2)$	– bivariate cumulative distribution function
$F_{\mathrm{T}}(y_1, y_2, y_3)$	– trivariate cumulative distribution function
$F_{\mathrm{M}}(\mathbf{y})$	– multivariate cumulative distribution function
S, $\widetilde{S}$, $S(y)$, $\widetilde{S}(y)$	– survival function
$S(y\vert\cdot)$	– conditional survival function
$S_{\mathrm{m}}(y)$	– mixture survival function
$S_{\mathrm{B}}(y_1, y_2)$	– bivariate survival function
$S_{\mathrm{M}}(\mathbf{y})$	– multivariate survival function
$H(y)$, $\widetilde{H}(y)$	– integrated intensity function
$H(y\vert\cdot)$	– conditional integrated intensity function
$h(y)$, $\widetilde{h}(y)$	– hazard, or intensity (function)
$h(y\vert\cdot)$	– conditional hazard or intensity (function)
$p(\mu)$	– probability density function for parameters
$p(\mu\vert\cdot)$	– conditional density for parameters

H_j	– subject's response history up to and including response j
ϕ	– Laplace transform
ϕ^{-1}	– inverse Laplace transform
$\phi(t_1,t_2)$	– bivariate Laplace transform
$\phi(\mathbf{t})$	– multivariate Laplace transform
u, $\tilde{u}$	– denotes (transformed) cdfs of y for the marginal distributions of a multivariate distribution
v, $\tilde{v}$	– denotes (transformed) survival functions of y for the marginal distributions of a multivariate distribution
$\breve{v}$	– denotes (transformed) integrated intensity functions of y for the marginal distributions of a multivariate distribution
$C_{\mathrm{B}}(u_1,u_2)$	– bivariate copula
$C_{\mathrm{T}}(u_1,u_2,u_3)$	– trivariate copula
$C_{\mathrm{F}}(u_1,u_2,u_3,u_4)$	– 4-variate copula
$C_{\mathrm{M}}(\mathbf{u})$	– multivariate copula
F	– denotes Fréchet classes
F_L	– Fréchet lower bound
F_U	– Fréchet upper bound
P	– product copula or independence
$\left.\begin{array}{l} C^*_{\mathrm{B}}(u_1,u_2), \\ C^{**}_{\mathrm{B}}(u_1,u_2), \\ C^{***}_{\mathrm{B}}(u_1,u_2) \end{array}\right\}$	– associated copulas in the bivariate case
$\widehat{C}_{\mathrm{B}}(v_1,v_2)$	– bivariate survival copula
$\widehat{C}_{\mathrm{M}}(\mathbf{v})$	– multivariate survival copula
$\overline{C}_{\mathrm{B}}(u_1,u_2)$	– bivariate joint survival function of univariate cdfs
$\overline{C}_{\mathrm{M}}(\mathbf{u})$	– multivariate joint survival function of univariate cdfs
ψ, ξ	– additive generators of the copula
$\psi^{[-1]}$, $\xi^{[-1]}$	– pseudo-inverses of additive generators
γ	– multiplicative generator of the copula
m	– number of margins
n	– number of observations or repeated measurements
t	– index for time
i, j, k	– other arbitrary indices

1
Ordinal models

Ordinal data have always been present in many different fields of research. The current worldwide trend to accumulate information has revealed the importance of categorical and more specifically ordinal responses. In parallel, the design of studies has become more complex in order to understand better the relationships among variables or factors influencing certain outcomes. This can for instance clearly be seen in the pharmaceutical sector where there is a growing concern for assessing patients' quality of life. In most cases, patients will be followed over time and have their state measured repeatedly. Hence, there is a growing need for elaborate analysis methods to take into account the dependence among observations recorded on the same individual.

Much work has been done for continuous responses. Although a number of distributions were available, the Gaussian distribution had the most success due to its very tractable mathematical properties. Due to the growth in computer power, various other methods have started to be investigated over the last decade. Unfortunately, few statisticians have been attracted to ordinal data leaving many topics to be investigated compared to other statistical fields.

This thesis will primarily be concerned with the analysis of dependent ordinal responses.

The first chapter introduces basic concepts including various regression models for univariate ordinal responses. Several different parametrizations are considered to take advantage of the order among categories.

The second chapter covers a variety of techniques used in different statistical fields to introduce dependence among observations. These concepts are the bases required to construct various multivariate distributions for ordinal data discussed in chapters four and five.

In chapter three, different particular combinations of distributions are considered in order to obtain closed form mixture distributions that will be useful for constructing dynamic longitudinal dependencies in chapter five.

Chapter four is devoted to models generated from copulas. Several specific types of dependencies are obtained by investigating different methods of combining Archimedean copulas. Most of these models will be most suitable for the analyses of clustered and multiple response data. Note that the term "multiple response" is used for datasets including several different kinds of responses observed per individual. On the other hand, the term "multivariate" is more general because it also includes repeated measurements of the same type of response.

The fifth chapter covers dynamic models for longitudinal data. This includes dynamically updated models and hidden Markov chains. Although these models are mainly oriented towards repeated measurements over time, dynamically updated models are in certain cases found to correspond to Archimedean copulas. When this occurs, such models could then also be used to analyze multiple response and clustered data.

Chapter six describes methods to represent graphically ordinal data collected over time as well as results from dynamic models.

Finally, chapter seven illustrates a few of these different methods through the analyses of clinical trials and smaller studies.

1.1 Variables

At the start of any study or clinical trial, the unit of observation must be defined. This can be any unit of interest, such as a person, an animal, a litter of mice, or a clinical centre. A group of these units or subjects can then be divided up according to differences in a common characteristic. For instance, this common characteristic might be sex, a measurement of blood pressure, or a personal appreciation of current health. This classifies the subjects into different categories depending on the type of characteristic they have. But such a classification will only be useful if all subjects have this characteristic and each subject can take only one type of this characteristic. Hence, a classification which creates such mutually exclusive and exhaustive categories will be called a variable. A variable can take different values, one corresponding to each category or type of characteristic. In the case of the variable sex, the possible values are male and female.

Because not all variables are the same, they can themselves be classified accordingly to the characteristics of their values.

- A variable is called a ratio variable when the values are numerical measures with a natural zero point.
- A variable is called an interval variable when the values are numerical measures without a natural zero point.
- A variable is integral if the values are counts.
- A variable is ordinal if the values are ordered but have no measure of distance between them.
- A variable is nominal if the values have no natural ordering and no numerical relationship among them.

The variables belonging to the two first types are commonly referred to as being continuous whereas the remaining ones are called discrete or categorical. A discrete or categorical variable that can only take two possible values is called a binary variable.

A variable can be moved downwards through this list but not upwards. Unfortunately this is not achieved without loss of information. For instance, measurements of blood pressure can be grouped into a small number of categories (discretized) or the order of an ordinal variable can always be ignored to obtain a nominal one. In a similar way, statistical methods are usually developed for variables of one particular type but could be applied to others above in this list. But it is always best to apply methods appropriate for the actual type of variable under consideration, which will then use as much as possible if not all the information available.

When variables are measured, two factors can influence the recording quality. The accuracy corresponds to how close the measurement made is to its “true” value whereas precision refers to the unit of measurement. For example, a watch always off by two minutes is not accurate and recording time of an event to the second is more precise than to half a minute. As all measuring instruments can only record to a finite precision, a discrete or categorical variable is always produced. But, in practice, the distinction between continuous and discrete variables is made accordingly to the number of categories they can take. Variables that can take a lot of values will generally be treated as continuous.

Another distinction still can be made among variables. A nominal variable is qualitative because its different categories refer to a characteristic or quality, not to a quantity. An integral or continuous variable is quantitative because the possible values it can take indicate a certain amount of the characteristic of interest. On the other hand, the position of an ordinal variable in this classification is not clear because the categories possess a magnitude of a characteristic but no actual measure of distance separating them.

An ordinal variable can often be thought to correspond to an underlying continuous variable which is usually not measurable. This is the approach that will be used in this text. In such a case, (if possible) the categories should be chosen carefully because it is often best to have them corresponding to similar ranges of the underlying continuous scale. This can enable one in some cases to take advantage of the variety of methods available for continuous variables.

Finally, a distinction is usually made between a response and a covariate (or explanatory variable). The response refers to the variable under investigation whereas covariates are additional variables collected to describe changes or explain patterns of this response. The main interest of this work is the analysis of ordinal responses with dependencies. All models developed for this purpose can include both continuous and discrete covariates.

1.2 Generalized linear models

Generalized linear models (Nelder and Wedderburn, 1972) are based on the exponential family of distributions which provides a range of univariate models. However, they are not meant to take into account the dependence among related observations. All members of the exponential family have their density or probability mass function of the form

$$f(y_i;\boldsymbol{\theta}_i) = a_0(\boldsymbol{\theta}_i)b(y_i)\mathrm{e}^{\mathbf{c}(y_i)\boldsymbol{\theta}_i} \tag{1.1}$$

where y_i is the response for individual i, $\boldsymbol{\theta}_i$ is the vector of canonical parameters, $a_0(\boldsymbol{\theta}_i)$ is the normalizing constant of the distribution, $\mathbf{c}(y_i)$ are the canonical statistics, and $b(y_i)$ is a remaining function of y. This exponential representation is called the canonical parametrization of the family. Though it is not obvious at first sight, this family includes the two most common discrete distributions. Indeed, the binomial (which includes the Bernoulli as a special case) and the Poisson distributions are part of this family; see respectively Equations (B.46) and (B.45).

Now, a regression model can be obtained by equating some function of the location parameter to a linear function of other parameters. This can then be written as

$$g(\mu_i) = \beta_0 + \sum_l \beta_l x_{il} \tag{1.2}$$

where β_0 (usually) is an unknown parameter for the intercept and β_l is (usually) an unknown parameter for the lth covariate x_{il} corresponding to individual i.

Notice that this is a function linking the distribution's location parameter μ_i to the regression. If $g(\mu_i) = \theta_i$, this function $g(\cdot)$ is called the canonical link. A linear model for the mean response is obtained when the identity link, $g(\mu_i) = \mu_i$, is chosen.

1.3 Univariate linear regression

Several methods are possible to analyze nominal and ordinal responses. The most common method is to classify the response values into a contingency table. Each cell of this table contains the number of observations recorded in a particular category, generally referred to as frequencies. Unfortunately, this is not most suitable when individual data are involved. Hence, an alternative approach will now be considered.

For simplicity, the special but important case of binary responses is first discussed. The two possible types of outcomes are usually referred to as "success" or "failure" and will respectively be represented by 1 and 0. It is generally assumed that the observations on different individuals occur independently. The response y_i for individual i can then be described by a Bernoulli distribution, see Equation (B.47). The Bernoulli probability mass function is

$$\begin{aligned} f(y_i;\pi_i) &= \pi_i^{y_i}(1-\pi_i)^{1-y_i} \\ &= (1-\pi_i)\left(\frac{\pi_i}{1-\pi_i}\right)^{y_i} \\ &= (1-\pi_i)\mathrm{e}^{y_i \ln\left[\frac{\pi_i}{1-\pi_i}\right]} \end{aligned} \tag{1.3}$$

where the (location or mean) parameter π_i denotes the probability of success for individual i, leaving $1-\pi_i$ as probability of failure (Agresti, 1990, p. 81). We see that $\ln\left[\frac{\pi_i}{1-\pi_i}\right]$ is the canonical link function.

This can now be extended to allow for nc categories (taking in some cases advantage of their order) as long as the constraint $\sum_{k=1}^{\texttt{nc}} \pi_{ik} = 1$ always holds. The remainder of this section covers different parametrizations of the distribution's location using this link function and then generalizations by the use of a variety of other link functions.

Note that the total number of categories and covariates are denoted respectively by nc and ncv and that all models collapses to the binary case if there are only two categories. Further, if there are more than two categories but no covariates included in the model then these parametrizations reduce to the multinomial model.

1.3.1 BINARY

The binary model (Agresti, 1990, pp. 91–97) is now straightforward with the canonical link function ($\texttt{nc} = 2$). The probability π_i of subject i being observed in category 1 is

$$\pi_i = \frac{\mathrm{e}^{\alpha + \sum_{l=1}^{\texttt{ncv}} \beta_l x_{il}}}{1 + \mathrm{e}^{\alpha + \sum_{l=1}^{\texttt{ncv}} \beta_l x_{il}}}$$

and $1 - \pi_i$ of being observed in category 0.

The probability of a sample can then be written as

$$\Pr(y_i; \pi_i) = \prod_i \left[(\pi_i)^{y_i} (1 - \pi_i)^{1 - y_i} \right]$$

which corresponds to a product of Bernoulli distributions.

Sometimes, several subjects will have the same combination of covariate values. In this case, a contingency table containing all different possible covariate patterns with non-empty cells can be built to improve computation time. The sample probability can then be written as

$$\Pr(y_i; \pi_i) = \prod_i \left[(\pi_i)^{n_{i1}} (1 - \pi_i)^{n_{i\bullet} - n_{i1}} \right]$$

where $n_{i\bullet}$ is the total number of subjects with covariate pattern i and n_{i1} is the number of subjects observed in category 1 with such a pattern.

The score equations and Fisher's information matrix are given in Appendix Section A.2 (Agresti, 1990, p. 118).

1.3.2 MULTINOMIAL

A first extension to allow additional categories, without taking their order into account, is called the multinomial model (Agresti, 1990, pp. 315–317). The probability π_{ik} of observation i being in response category k is

$$\pi_{ik} = \frac{\mathrm{e}^{\alpha_k + \sum_{l=1}^{\texttt{ncv}} \beta_{kl} x_{il}}}{1 + \sum_{h=1}^{\texttt{nc}-1} \mathrm{e}^{\alpha_h + \sum_{l=1}^{\texttt{ncv}} \beta_{hl} x_{il}}}$$

for all but the last category which is $1 - \sum_{h=1}^{\texttt{nc}-1} \pi_{ih}$ (equivalent to $\alpha_{\texttt{nc}} = \beta_{\texttt{nc}l} = 0, \forall\, l$).

The sample probability can then be written as

$$\Pr(y_i;\pi_i) = \prod_i \left[\prod_{h=1}^{\mathtt{nc}-1} (\pi_{ih})^{n_{ih}} \left(1 - \sum_{h=1}^{\mathtt{nc}-1} \pi_{ih} \right)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \right]$$

The score equations and Fisher's information matrix are given in Appendix Section A.3.

1.3.3 SIMPLIFIED MULTINOMIAL

A simplification of the previous model can be obtained by constraining the covariates to have the same coefficients for each possible response category with only the intercept changing. The probability π_{ik} of observation i being in response category k is

$$\pi_{ik} = \frac{\mathrm{e}^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}{1 + \sum_{h=1}^{\mathtt{nc}-1} \mathrm{e}^{\alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}$$

for all but the last category which is $1 - \sum_{h=1}^{\mathtt{nc}-1} \pi_{ih}$.

The sample probability can then be written as

$$\Pr(y_i;\pi_i) = \prod_i \left[\prod_{h=1}^{\mathtt{nc}-1} (\pi_{ih})^{n_{ih}} \left(1 - \sum_{h=1}^{\mathtt{nc}-1} \pi_{ih} \right)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \right]$$

The score equations and Fisher's information matrix are given in Appendix Section A.4.

1.3.4 CONTINUATION RATIO

A second extension to allow additional categories, this time taking their order taken into account, is called the continuation ratio model (Agresti, 1990, pp. 319–320; Greenland, 1994). This model contrasts all categories on one side of the scale to the immediately following one. It is therefore not symmetric, allowing both upward and downward parametrizations. Note that only one of these should be used, depending on the question to be answered.

- **Upwards**
  If p_{ik} is defined as

  $$p_{ik} = \frac{\mathrm{e}^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}{1 + \mathrm{e}^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}} \tag{1.4}$$

  for all but the last category then the probability π_{ik} of observation i being in response category k is

$$
\begin{aligned}
\pi_{i1} &= p_{i1}p_{i2}p_{i3}\dots p_{i,\mathtt{nc}-1}\\
\pi_{i2} &= (1-p_{i1})p_{i2}p_{i3}\dots p_{i,\mathtt{nc}-1}\\
\pi_{i3} &= (1-p_{i2})p_{i3}\dots p_{i,\mathtt{nc}-1}\\
&\vdots\\
\pi_{i,\mathtt{nc}} &= (1-p_{i,\mathtt{nc}-1})
\end{aligned}
$$

The sample probability can then be written as

$$
\Pr(y_i;\pi_i) = \prod_i \left[(p_{i1})^{n_{i1}}(1-p_{i1})^{n_{i2}}\dots(p_{i,\mathtt{nc}-1})^{\sum_{h=1}^{\mathtt{nc}-1} n_{ih}}(1-p_{i,\mathtt{nc}-1})^{n_{i\bullet}-\sum_{h=1}^{\mathtt{nc}-1} n_{ih}}\right]
$$

- **Downwards**
  If p_{ik} is defined as

$$
p_{ik} = \frac{\mathrm{e}^{\alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}}}{1+\mathrm{e}^{\alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}}} \tag{1.5}
$$

  for all but the first category then the probability π_{ik} of observation i being in response category k is

$$
\begin{aligned}
\pi_{i1} &= (1-p_{i2})\\
\pi_{i2} &= p_{i2}(1-p_{i3})\\
\pi_{i3} &= p_{i2}p_{i3}(1-p_{i4})\\
&\vdots\\
\pi_{i,\mathtt{nc}} &= p_{i2}p_{i3}\dots p_{i,\mathtt{nc}}
\end{aligned}
$$

  The sample probability can then be written as

$$
\Pr(y_i;\pi_i) = \prod_i \left[(1-p_{i2})^{n_{i\bullet}-\sum_{h=2}^{\mathtt{nc}} n_{ih}}(p_{i2})^{\sum_{h=2}^{\mathtt{nc}} n_{ih}}\dots(1-p_{i,\mathtt{nc}})^{n_{i,\mathtt{nc}-1}}(p_{i,\mathtt{nc}})^{n_{i,\mathtt{nc}}}\right]
$$

The score equations and Fisher's information matrix are given in Appendix Section A.5.

1.3.5 PROPORTIONAL ODDS

Another extension to allow additional categories, with their order taken into account, is called the proportional odds model (McCullagh, 1980; Agresti, 1990, pp. 322–324). This model contrasts lower to higher categories around various cut points and is therefore symmetric.

If p_{ik} is defined as

$$p_{ik} = \frac{e^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}{1 + e^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}} \tag{1.6}$$

for all but the last category then the probability π_{ik} of observation i being in response category k is

$$\begin{aligned}
\pi_{i1} &= p_{i1} \\
\pi_{i2} &= p_{i2} - p_{i1} \\
&\vdots \\
\pi_{i,\mathtt{nc}} &= 1 - p_{i,\mathtt{nc}-1}
\end{aligned}$$

The sample probability can then be written as

$$\Pr(y_i; \pi_i) = \prod_i [(p_{i1})^{n_{i1}} (p_{i2} - p_{i1})^{n_{i2}} \ldots (1 - p_{i,\mathtt{nc}-1})^{n_{i,\mathtt{nc}}}]$$

The score equations and Fisher's information matrix are given in Appendix Section A.6.

1.3.6 ADJACENT CATEGORIES

The last extension considered to allow additional categories, with their order taken into account, is called the adjacent categories model(Agresti, 1990, p. 318). This model contrasts immediately adjacent categories and is therefore symmetric.

The probability π_{ik} of observation i being in response category k is

$$\pi_{ik} = \frac{e^{-\sum_{g=1}^{k} \alpha_g + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}{1 + \sum_{h=1}^{\mathtt{nc}-1} e^{-\sum_{g=1}^{h} \alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}}}$$

for all but the last category which is $1 - \sum_{h=1}^{\mathtt{nc}-1} \pi_{ih}$.

The sample probability can then be written as

$$\Pr(y_i; \pi_i) = \prod_i \left[\prod_{h=1}^{\mathtt{nc}-1} (\pi_{ih})^{n_{ih}} \left(1 - \sum_{h=1}^{\mathtt{nc}-1} \pi_{ih} \right)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \right]$$

The score equations and Fisher's information matrix are given in Appendix Section A.7.

1.3.7 LINKS

As can be seen in Equation (1.3), the Bernoulli distribution does indeed have the characteristic form of the exponential family shown in Equation (1.1). It has

$\ln\left[\frac{\pi_i}{1-\pi_i}\right]$ as canonical link which is called the logit link. Hence, with this link, the distribution's location is described by the S-shaped curve

$$\pi_i = \frac{\mathrm{e}^{\alpha+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}}}{1+\mathrm{e}^{\alpha+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}}} = \frac{1}{1+\mathrm{e}^{-\left(\alpha+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}\right)}} \tag{1.7}$$

called the logistic regression function due to its similarity to the logistic cumulative distribution function (cdf) presented in Equation (B.29). The parameter α is the intercept and represents the distribution's location when all covariates are set to zero. Note that the latter form is computationally more attractive as it is faster to compute.

This suggests that the binary model can be generalized by allowing the distribution's location to be described by

$$\mu_i = F\left(\alpha+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}\right)$$

where $F(\cdot)$ is a continuous cdf. In this case, the canonical link presented in Equation (1.2) would be replaced by the corresponding quantile function, $Q(\cdot)$.

$$Q(\mu_i) = \alpha+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}$$

McCullagh (1980) shows that this generalization can also be applied to proportional odds models. This can be seen from Equation (1.6) which also corresponds to the logistic cdf presented in Equation (B.29).

$$Q(\mu_i) = \alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il} \tag{1.8}$$

Two link functions are suggested in that paper. The probit link

$$\Phi^{-1}(\mu_i) = \alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}$$

corresponding to the Gaussian distribution, where $\Phi^{-1}(\cdot)$ is defined in Equation (B.43), and the complementary log-log

$$\ln[-\ln(1-\mu_i)] = \alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il} \tag{1.9}$$

corresponding to the Gompertz distribution, see Equation (B.19).

McCullagh (1980) also relates models with a complementary log-log link to proportional hazards models (Cox, 1972). This is done by rewriting Equation (1.9) as

$$\begin{aligned} -\ln(1-\mu_i) &= \mathrm{e}^{\alpha_k+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}} \\ &= \theta_k \mathrm{e}^{-\sum_{l=1}^{\mathtt{ncv}}\phi_l x_{il}} \end{aligned} \tag{1.10}$$

where $\theta_k = \mathrm{e}^{\alpha_k}$ and $\phi = -\beta$.

A related link function is the log-log link

$$\ln[-\ln(\mu_i)] = \alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \tag{1.11}$$

corresponding to the Gumbel distribution, see Equation (B.17). Indeed, if the complementary log-log link models the probability of a success then the probability of a failure will be modelled by the log-log link (Agresti, 1990, p. 105).

For the two continuation ratio parametrizations, it can be seen that Equations (1.4) and (1.5) also corresponds to the logistic cdf presented in Equation (B.29). The link function is identical to one used for the proportional odds models, see Equation (1.8).

Although generalization is possible, it has not often been implemented. This can be explained by the additional complexity in estimating the parameters when other links than the logit are used. Fortunately, this is no longer a problem as numerical optimization can be used instead (see Subsection 1.4.2).

Finally, many other quantile functions could be used as links possibly incorporating additional unknown parameters as is suggested in the discussion following the paper of McCullagh (1980). For instance, quantile functions of generalized distributions could be useful in selecting among several known distributions. But any model with such a link function will no longer be linear in its parameters. Hence, numerical optimization will have to be used in order to obtain the maximum likelihood estimates of all parameters (see Subsection 1.4.2).

1.4 Maximum likelihood estimates

1.4.1 LINEAR REGRESSION

The likelihood corresponding to a particular set of independent observations is obtained by multiplying together the density or probability mass function for each of these observations.

$$\mathrm{L}(y_i;\mu_i) = \prod_i f(y_i;\mu_i)$$

For linear regression, parameter estimates are obtained by solving the first derivatives set to zero called the score equations. This is done along with the second derivatives called Fisher's information matrix if using Newton-Raphson method (Agresti, 1990, pp. 114–117). The solutions obtained are called the maximum likelihood estimates. Note that by "linear regression", it is implied that the model is linear in its unknown parameters, not in possible covariates included in the regression model.

If there is linearity in the parameters to be estimated, only a small number of specific derivatives written in a general form are required for each model. For the first derivatives, two equations are usually sufficient. One is obtained by taking a derivative with respect to an intercept parameter and the other with respect to a covariate parameter. However, it is slightly more complex to obtain the second derivatives for a nominal or ordinal parametrization (nc ≥ 3). A minimum of three equations are required. These are obtained with respect to two intercept parameters, to two covariate parameters, and to an intercept and a covariate parameter. But in most cases, for ordinal data additional (second derivative) equations are required due to special cases such as derivatives for the first and last categories or second derivatives taken twice with respect to the same parameter. Because Fisher's information matrix is symmetric, the set of second derivative equations is obtained regardless of the order in which the two derivatives are performed (under the appropriate regularity conditions).

Each of the equations obtained for a particular model is written in a general form so that maximum likelihood estimates can be obtained for any linear regression by repeating the different score equations to match the number of categories and covariates included as well as extending accordingly Fisher's information matrix.

1.4.2 NON-LINEAR REGRESSION

When the regression is non-linear, maximum likelihood estimates for the parameters can still be obtained by using the set of first and second derivatives. Unfortunately, these can no longer be written in a general form as they will be specific to each particular non-linear regression chosen. Hence, the score equations and Fisher's information matrix given in Appendix A cannot be used in this case. Fortunately, nowadays computers are fast enough to be able to optimize numerically most likelihoods in a reasonable amount of time.

In order to fit non-linear models, specialized software is required. For instance, link functions with additional parameters or cdfs such as the ones found in Appendix B must be programmed. But most of all, a special interface is required to specify non-linear regressions.

In a linear regression formula, parameters can be omitted as their position can be implicitly assumed. Unfortunately, this assumption is no longer possible for any non-linear regression. Hence, a notation is required where the parameters appear explicitly. This can be obtained by adapting the Wilkinson and Rogers (1973) notation. To implement this improved notation, the software must be able to distinguish automatically among existing variables, unknown parameters, and mathematical functions included in the regression formula. Such an interface has been implemented in the freely available software R, a fast clone of S-Plus developed by Ross Ihaka, Robert Gentleman, and the R core group (Ihaka and Gentleman, 1996).

Finally, additional shape or family parameter(s) introduced through link functions could also depend (non-)linearly on covariates. For instance, a distribution could perhaps be changing shape over time. Hence, the nonlinear interface can also be used for such parameters in order to obtain models as flexible as possible.

1.5 Previous work on multivariate ordinal models

The literature on univariate ordinal models is vast; no attempt will be made to cover it here. Due to the generalization brought by McCullagh (1980) and to the wide availability of software, the proportional odds model has since been the most widely used parametrization for univariate ordinal data. However, Greenland (1994) provides an alternative choice of models and discusses different situations motivating the use of several of these particular parametrization.

Another, quite different, approach to ordinal data has been proposed by Goodman (1979 and 1981). It will not be further considered here.

For a general introduction to ordinal data, see Agresti (1990).

Since, several extensions have been considered in order to handle multivariate ordinal responses. The most popular approach consists in including random effects. Bartholomew (1980) discusses multiple integration of the random effects in general, for both binary and ordinal responses. The particular case of a probit link is considered for ordinal responses because the model can then always be expressed in terms of bivariate integrals. Jansen (1990) uses Gaussian quadrature to approximate the multiple integrals and applies the EM algorithm to obtain the parameter estimates. Ezzet and Whitehead (1991) discuss random effects for two period cross-over trials. In this case, the bivariate random effect model is evaluated numerically using the Newton-Raphson method allowing choice of different distributions for the random effects. Hedeker and Gibbons (1994) provide an implementation of multiple random effects for both an ordinal probit and logit model. For this, Gauss-Hermite quadrature is used to integrate out numerically the multivariate Gaussian random effects. Multilevel Gaussian random effects for ordinal data have also been considered by Keen and Engel (1997) — using iterative re-weighted restricted maximum likelihood to obtain the parameter estimates — and Ribaudo *et al.* (1999) — using second order penalized quasi-likelihood estimation. The adjacent category and proportional odds parametrizations are used by Hartzel *et al.* (2001) to model a multicentre ordinal clinical trial. A multivariate Gaussian random effect is once again chosen where the integrals are approximated by adaptive Gauss-Hermite quadrature and the parameter estimates are obtained using the Newton-Raphson or quasi-Newton methods.

Due to the complexity of evaluating multiple integrals, other random effect models have developed where the multivariate distribution could be obtained in closed form. This general approach is based on the mixture model proposed by Farewell (1982) to capture overdispersion in ordinal data. The mixture proposed combines an extreme value distribution, corresponding to either the complementary log-log, with a log-gamma mixing distribution. Conaway (1990) uses a multivariate generalized logistic distribution to model binary data. Crouchley (1995) extends this model to ordinal responses and generalizes the mixing distribution to the power variance family (Hougaard, 1986a). The multivariate model obtained includes the gamma, inverse Gaussian, and positive stable distributions as special cases. Ten Have (1996) and Pulkstenis *et al.* (1998) adapted the model based on the log-gamma mixing distribution to include discrete failure times with ordinal responses. In the later case, this is used to model dropouts.

The Rasch model can be interpreted as a non-parametric random effect (Lindsay *et al.*, 1991). This model was first extended for multinomial responses by Conaway (1989) and then for ordinal responses by Agresti and Lang (1993). Kenward and Jones (1991) applied the multinomial Rasch model to cross-over trials.

Other types of dependencies have also been introduced for longitudinal studies. Generalized estimating equations which replace the multinomial dispersion matrix have been considered by Gange *et al.* (1995). Lindsey *et al.* (1997) induce serial correlation among repeated ordinal observation by including the lagged response as a covariate in a continuation ratio model. First order Markov chain models are developed by Albert *et al.* (1997) for ordinal responses with a proportional odds parametrizations.

The Dale model and its multivariate extension, the Plackett model, have both been used to analyze ordinal data. Molenberghs and Lesaffre (1994) apply these models to cross-over trials and longitudinal clinical trials. Kenward *et al.* (1994) consider the application of such models using maximum likelihood and generalized estimating equations for the analysis of longitudinal studies. Lesaffre *et al.* (1996) and Molenberghs *et al.* (1997) use these models to analyze longitudinal ordinal studies taking the dropout mechanism in account. Lapp *et al.* (1998) consider these models in the case of multiple responses.

Finally, Agresti (1989 and 1999) and Gibbons and Hedeker (2000) provide reviews the different models available to analyze multivariate ordinal data.

2
Inducing dependencies

Many different types of associations or dependencies can occur among observations. Although fundamental differences reside among these, the proposed approach to introduce certain dependency structures will be based on a common theoretical background gathered from several statistical fields.

The first section introduces mixtures including the particular case of Laplace transforms. Because Laplace transforms can often be obtained in a closed form, they will extensively be used in the remaining chapters. Thus, several properties are investigated as well as some generalization methods.

The second section considers three possible methods to construct multivariate distributions. In the first approach, multivariate distribution are obtained from products of conditional distributions. On the other hand, the second studies multivariate dependence structures which are independent from their univariate marginal distributions. These are called copulas and can be produced for instance by combining strictly decreasing or increasing convex functions. In this particular case, the Archimedean copulas are obtained which are investigated further due to their tractable properties. Finally in the last approach, Laplace transforms are extended to the multivariate case.

2.1 Mixtures

A first approach consists in introducing an additional parameter into an available distribution. This parameter will then be used to induce the desired dependence structure in Chapter 5. One possible procedure to achieve this consists in combining two known distributions. Such a resulting distribution is usually called a compound or mixture distribution.

DEFINITION 2.1 *[Wetherill, 1961; Harvey, 1989, p. 163]*

A mixture density, $f_{\mathrm{m}}(y)$, is obtained by integration of the product of a density and a conditional density.

$$f_{\mathrm{m}}(y) = \int_0^\infty f(y|\lambda)f(\lambda)\mathrm{d}\lambda \tag{2.1}$$

The density $f(\lambda)$ is called the *mixing distribution* and the conditional density $f(y|\lambda)$ is a density, given the random parameter of this mixing distribution. This can be interpreted as the parameter λ of $f(y|\lambda)$ varying randomly in a population according to $f(\lambda)$. □

Under regularity conditions, Equation (2.1) can be rewritten in terms of cdfs

$$\begin{aligned} F_{\mathrm{m}}(y) &= \int_0^\infty F(y|\lambda) f(\lambda) \mathrm{d}\lambda \\ &= \int_0^\infty F(y|\lambda) \mathrm{d}F(\lambda) \end{aligned} \tag{2.2}$$

where $F_{\mathrm{m}}(y)$ is the resulting mixture cumulative distribution function, $F(y|\lambda)$ is the conditional cumulative distribution function, and $F(\lambda)$ is the mixing cumulative distribution function (Johnson and Kotz, 1969, pp. 27–28).

A similar relationship can also be obtained in terms of survival functions

$$\begin{aligned} S_{\mathrm{m}}(y) &= 1 - F_{\mathrm{m}}(y) \\ &= 1 - \int_0^\infty F(y|\lambda) \mathrm{d}F(\lambda) \\ &= \int_0^\infty \mathrm{d}F(\lambda) - \int_0^\infty F(y|\lambda) \mathrm{d}F(\lambda) \\ &= \int_0^\infty [1 - F(y|\lambda)] \mathrm{d}F(\lambda) \\ &= \int_0^\infty S(y|\lambda) \mathrm{d}F(\lambda) \end{aligned} \tag{2.3}$$

where $S_{\mathrm{m}}(y)$ is the resulting mixture survival function, and $S(y|\lambda)$ is the conditional survival function.

A mixture will only have additional parameters if at least one parameter from the mixing distribution and one from the conditional density remain unconfounded. In such a case, this can result in the creation of a new distribution.

2.1.1 LAPLACE TRANSFORMS

A tractable form of mixtures, using a Laplace transform, arises when the exponential distribution is inserted as the conditional survival function of Equation (2.3).

DEFINITION 2.2 *[Abramowitz and Stegun, 1965, p. 1020; Hougaard, 1984; Joe, 1997, pp. 85–86 and 119; Nelsen, 1999, p. 65]*

Let $F(\lambda)$ be a univariate cdf of a positive random variable with $F(0) = 0$ and $\phi(y)$ be the Laplace transform of $F(\lambda)$. Then

$$\phi(y) = \int_0^\infty \mathrm{e}^{-y\lambda} \mathrm{d}F(\lambda) \qquad y > 0 \tag{2.4}$$

is an exponential mixture survival function, if $F(\lambda)$ is a mixing cdf. Note that $\phi(-y)$ is the moment generating function of $F(\lambda)$. □

The density corresponding to a Laplace transform can be obtained by inserting the exponential distribution into Equation (2.1),

$$f_{\mathrm{m}}(y) = \int_0^\infty \lambda \mathrm{e}^{-y\lambda} f(\lambda) \mathrm{d}\lambda = \frac{\partial[1 - \phi(y)]}{\partial y} = -\phi'(y)$$

where $f_{\mathrm{m}}(y)$ is an exponential mixture distribution (Cox and Oakes, 1984, p. 19). This is the Laplace transform of $\lambda f(\lambda)$.

PROPERTY 2.3 *[Abramowitz and Stegun, 1965, p. 1020; Salas and Hille, 1990, pp. 216–217; Joe, 1997, p. 373]*

Laplace transforms are strictly decreasing functions because the derivative

$$\phi^{'}(y) = -\int_0^\infty \mathrm{e}^{-y\lambda}\lambda f(\lambda)\mathrm{d}\lambda$$

is negative. Under regularity conditions, higher order derivatives can be obtained by

$$\phi^{(m)}(y) = (-1)^m \int_0^\infty \mathrm{e}^{-y\lambda}\lambda^m f(\lambda)\mathrm{d}\lambda$$

which leads to the conclusion that a Laplace transform is also a convex function because its second derivative is positive. ○

Further useful relationships can now be obtained from survival theory (Cox and Oakes, 1984, pp. 14–15)

$$f(y|\lambda) = h(y|\lambda)\mathrm{e}^{-H(y|\lambda)} \tag{2.5}$$
$$S(y|\lambda) = 1 - F(y|\lambda) \tag{2.6}$$
$$H(y|\lambda) = -\ln(S(y|\lambda))$$
$$h(y|\lambda) = \frac{f(y|\lambda)}{S(y|\lambda)}$$

where $f(y|\lambda)$ is a conditional density, $S(y|\lambda)$ is the corresponding conditional survival function, $H(y|\lambda)$ is the corresponding conditional integrated hazard, and $h(y|\lambda)$ is the corresponding conditional hazard or intensity.

In the case of the exponential distribution, these relationships are

$$f(y|\lambda) = \lambda\mathrm{e}^{-\lambda y} \tag{2.7}$$
$$S(y|\lambda) = \mathrm{e}^{-\lambda y}$$
$$H(y|\lambda) = \lambda y$$
$$h(y|\lambda) = \lambda$$

(Cox and Oakes, 1984, p. 16; others are given in Appendix B). Two important properties of the exponential distribution can now clearly be seen. It has a constant hazard and the conditional parameter λ has a multiplicative effect in the integrated hazard.

A generalization can be obtained by allowing the function argument y to be replaced by an arbitrary function of y, say $G(y)$, in Equation (2.5)

$$f(G(y)|\lambda) = h(G(y)|\lambda)\mathrm{e}^{-H(G(y)|\lambda)}\frac{\partial G(y)}{\partial y}$$

where $\frac{\partial G(y)}{\partial y}$ is the Jacobian corresponding to the transformation imposed by the function $G(y)$.

This can for example be applied to the exponential density in Equation (2.7).

$$f(G(y)|\lambda) = \lambda \mathrm{e}^{-\lambda G(y)} \frac{\partial G(y)}{\partial y} \tag{2.8}$$

Letting the function $G(y)$ be an arbitrary integrated hazard, say $\widetilde{H}(y)$, Equation (2.8) becomes

$$\begin{aligned} f(\widetilde{H}(y)|\lambda) &= \lambda \mathrm{e}^{-\lambda \widetilde{H}(y)} \frac{\partial \widetilde{H}(y)}{\partial y} \\ &= \lambda \mathrm{e}^{-\lambda \widetilde{H}(y)} \widetilde{h}(y) \end{aligned} \tag{2.9}$$

where $\widetilde{h}(y)$ is the hazard corresponding to the integrated hazard $\widetilde{H}(y)$.

The corresponding conditional survival function, integrated hazard, and hazard to the conditional density $f(\widetilde{H}(y)|\lambda)$ are

$$S(\widetilde{H}(y)|\lambda) = \mathrm{e}^{-\lambda \widetilde{H}(y)} \tag{2.10}$$

$$H(\widetilde{H}(y)|\lambda) = \lambda \widetilde{H}(y) \tag{2.11}$$

$$h(\widetilde{H}(y)|\lambda) = \lambda \widetilde{h}(y)$$

Equation (2.11) corresponds to imposing a multiplicative effect of the conditional parameter λ in the integrated hazard as was suggested by Hougaard (1984).

Simplification of Equation (2.10) results in the power of a survival function.

$$\begin{aligned} S(\widetilde{H}(y)|\lambda) &= \mathrm{e}^{-\lambda \widetilde{H}(y)} \\ &= \mathrm{e}^{-\lambda[-\ln(\widetilde{S}(y))]} \\ &= \widetilde{S}^{\lambda}(y) \end{aligned}$$

This relationship can now be used as the conditional survival function of a mixture distribution to obtain a special case of Laplace transform.

THEOREM 2.4 *[Joe, 1997, p. 86]*

For an arbitrary survival function $S_{\mathrm{m}}(y)$ of a strictly positive random variable (no mass at zero), there exists a unique survival function $\widetilde{S}(y)$ such that

$$\begin{aligned} \phi(\widetilde{H}(y)) &= \int_0^\infty \mathrm{e}^{-\lambda \widetilde{H}(y)} \mathrm{d}F(\lambda) \\ &= \int_0^\infty \widetilde{S}^{\lambda}(y) \mathrm{d}F(\lambda) \\ &= S_{\mathrm{m}}(y) \end{aligned} \tag{2.12}$$

where $\widetilde{S}^{\lambda}(y)$ is the power of a survival function and $\widetilde{H}(y)$ is the corresponding integrated hazard. This Laplace transform is indeed a survival function because the transformation $\widetilde{H}(y)$ applied to the domain of ϕ does not affect its range. This theorem holds for any survival function $S_{\mathrm{m}}(y)$ because it is invertible

$$\widetilde{S}(y) = \mathrm{e}^{-\phi^{-1}(S_{\mathrm{m}}(y))} \tag{2.13}$$

where ϕ^{-1} is an inverse Laplace transform. □

A similar relationship also exists for cdfs when the function $G(y)$ is chosen to be $-\ln[\widetilde{F}(y)]$ where $\widetilde{F}(y)$ is an arbitrary cdf (Joe, 1997, p. 86).

LEMMA 2.5 *[Joe, 1997, p. 86]*

For an arbitrary cdf $F_{\mathrm{m}}(y)$, there exists a unique cdf $\widetilde{F}(y)$ such that

$$\begin{aligned}\phi(-\ln[\widetilde{F}(y)]) &= \int_0^\infty \widetilde{F}^{\lambda}(y)\mathrm{d}F(\lambda)\\ &= F_{\mathrm{m}}(y)\end{aligned} \tag{2.14}$$

where $\widetilde{F}^{\lambda}(y)$ is the power of a cdf. In this case, the result is a cdf because the transformation $-\ln[\widetilde{F}(y)]$ applied to the domain of ϕ inverts its range. Similarly, this lemma holds for any cdf $F_{\mathrm{m}}(y)$ because it is invertible

$$\widetilde{F}(y) = \mathrm{e}^{-\phi^{-1}(F_{\mathrm{m}}(y))} \tag{2.15}$$

where ϕ^{-1} is an inverse Laplace transform. □

Theorem 2.4 and Lemma 2.5 are the foundation blocks for mixtures of powers used in Sections 2.2.3 and 4.3.4.

2.2 Multivariate distributions

Many different techniques can be used to obtain multivariate distributions. But first the concept of a multivariate distribution must be defined. As the responses under study are nominal or ordinal, the following definition is given in terms of multivariate cdfs.

DEFINITION 2.6 *[Schweizer and Sklar, 1983, p. 82; Nelsen, 1999, pp. 40–41]*

Let m be an integer greater than or equal to two. An m-dimensional cdf is a function F_{M} satisfying the conditions

(1) $\mathrm{Dom}(F_{\mathrm{M}}) = \mathbf{R}^m$ where $\mathrm{Dom}(F_{\mathrm{M}})$ denotes the domain of F_{M}.
(2) The range of F_{M} includes zero (grounded).
(3) The volume of F_{M} is non-negative for any m-dimensional interval (m-increasing).
(4) $F_{\mathrm{M}}(\infty,\infty,\ldots,\infty) = 1$

A joint or multivariate cdf is a function F_{M} for which there is an integer $m \geq 2$ such that F_{M} is an m-dimensional cdf. □

From this definition, a general approach to obtain multivariate distributions is to consider the necessary and sufficient conditions for a right continuous function G on $\mathbf{R}^m$ to be a multivariate distribution. These are that G has the appropriate domain and that the rectangle inequality holds. If G has mth-order derivatives, this last condition is fulfilled when these m partial derivatives have the appropriate signs. Note that using this replacement condition will also ensure that the multivariate distribution has a closed form density.

For example, a multivariate cdf (F_{M}) is obtained if $0 \leq F_{\mathrm{M}} \leq 1$ and all m partial derivatives of F_{M} are positive (Joe, 1997, pp. 11–12). On the other hand, a multivariate survival function (S_{M}) is obtained if $1 \geq S_{\mathrm{M}} \geq 0$ and the m partial derivatives of S_{M} have alternating signs (Nelsen, 1999, p. 122; Hougaard, 2000, p. 222).

Three approaches which fulfill these two conditions will now be considered. Note that, in Subsections 2.2.2 and 2.2.3, multivariate distributions are obtained from given marginal distributions. In this case, the marginal distributions will usually be either (transformed) cdfs of y denoted by u or (transformed) survival functions of y denoted by v.

2.2.1 PRODUCTS OF CONDITIONAL DISTRIBUTIONS

A first approach is to consider the relationship between conditional and multivariate distributions.

$$f_j(y_j|y_1,\ldots,y_{j-1}) = \frac{f_{\mathrm{M}}(y_1,\ldots,y_j)}{f_{\mathrm{M}}(y_1,\ldots,y_{j-1})} \tag{2.16}$$

Note that in general, the ratio of two multivariate distributions does not produce a conditional distribution of the same form.

From recursion of Equation (2.16), a multivariate distribution can be constructed from a hierarchical series of conditional distributions.

$$f_{\mathrm{M}}(y_1,\ldots,y_m) = f_1(y_1)f_2(y_2|y_1)\ldots f_m(y_m|y_1,\ldots,y_{m-1}) \tag{2.17}$$

The multivariate distribution obtained is a product of terms implying that the conditional distributions are independent. However, it is important to notice that there is a strict one-way ordering among the conditional distributions. Hence, care must be taken because each response conditions on a different number of previous responses.

Such an ordering is suitable when there is a specific ordering among the series of observations such as with time. But usually responses do not depend on all of the information brought by previous observations. Hence, dependence of a response on previous ones can be chosen in order to attain a stationary process. A common method used is to create a Markov process of order M where responses $m-k$ up to m are assumed to be (conditionally) independent for $k > M$. Note that the first terms remain special because they condition on different numbers of previous observations. This requires some initial conditions, at the start of the series, to be estimated or specified.

Many other types of dependencies can also be created. In some cases, the multivariate distribution obtained may then be invariant to reordering. Such dependencies are therefore suitable to model clustered or multiple response data.

Hence, the product of conditional distributions is a very flexible method to create a multivariate distribution. Indeed, many different multivariate distributions can be obtained. This is due to all the different combinations of series of conditional distributions and to the various specific types of dependencies which can be induced.

2.2.2 COPULAS

A multivariate distribution can also be obtained by defining separately the univariate marginal distributions and the dependence structure linking them together. But first several definitions are required to be in a position to state this relationship.

DEFINITION 2.7 *[Schweizer and Sklar, 1983, pp. 82–83; Nelsen, 1999, p. 8]*
A function $\widetilde{C}_{\mathrm{B}}$ that satisfies the conditions

(1) Dom $\left(\widetilde{C}_{\mathrm{B}}\right) = A_1 \times A_2$ where A_1 and A_2 are subsets of the interval $[0,1]$ containing 0 and 1.
(2) The range of $\widetilde{C}_{\mathrm{B}}$ includes zero (grounded).
(3) The volume of $\widetilde{C}_{\mathrm{B}}$ is non-negative for any two-dimensional interval (2-increasing).
(4) $\widetilde{C}_{\mathrm{B}}(u_1,1) = u_1$ and $\widetilde{C}_{\mathrm{B}}(1,u_2) = u_2$ for every $u_1 \in A_1$ and every $u_2 \in A_2$.

is called a two-dimensional subcopula. □

A subclass of main interest contains subcopulas whose domain is the entire unit surface $[0,1]^2$.

DEFINITION 2.8 *[Frank, 1979; Genest and MacKay, 1986b; Nelsen, 1999, p. 8]*
A function C_{B} mapping $[0,1]^2$ to $[0,1]$ satisfying the conditions

(1) $C_{\mathrm{B}}(u_1,1) = u_1$ and $C_{\mathrm{B}}(1,u_2) = u_2$ for $u_1,u_2 \in [0,1]$
(2) $C_{\mathrm{B}}(u_1,0) = C_{\mathrm{B}}(0,u_2) = 0$
(3) The range of C_{B} includes zero (grounded).
(4) The volume of C_{B} is non-negative for any two-dimensional interval (2-increasing). Hence, for every $u_{1a},u_{1b},u_{2a},u_{2b}$ in $[0,1]$ such that $u_{1a} \leq u_{1b}$ and $u_{2a} \leq u_{2b}$,

$$C_{\mathrm{B}}(u_{1b},u_{2b}) - C_{\mathrm{B}}(u_{1b},u_{2a}) - C_{\mathrm{B}}(u_{1a},u_{2b}) + C_{\mathrm{B}}(u_{1a},u_{2a}) \geq 0$$

is called a bivariate copula. □

These definitions can now be extended to the multivariate case.

DEFINITION 2.9 *[Schweizer and Sklar, 1983, pp. 82–83; Nelsen, 1999, pp. 39–40]*
Let m be an integer greater or equal to two. A function $\widetilde{C}_{\mathrm{M}}$ that satisfies the conditions

(1) Dom $\left(\widetilde{C}_{\mathrm{M}}\right) = A_1 \times A_2 \times \ldots \times A_m$ where all A_j are subsets of the interval $[0,1]$ containing 0 and 1, for $j = 1,\ldots,m$.
(2) The range of $\widetilde{C}_{\mathrm{M}}$ includes zero (grounded).
(3) The volume of $\widetilde{C}_{\mathrm{M}}$ is non-negative for any m-dimensional interval (m-increasing).

(4) $\widetilde{C}_{\mathrm{M}}$ has (one-dimensional) margins $\widetilde{C}_j$ which satisfy $\widetilde{C}_j(u) = u$ for all $u \in A_j$, for $j = 1, \ldots, m$.

is called an m-dimensional subcopula. □

Thus, a multivariate distribution with all univariate margins being $U(0,1)$ is a function C_{M} called a (multivariate) copula.

DEFINITION 2.10 *[Schweizer and Sklar, 1983, pp. 78–85; Molenberghs, 1992, pp. 59–60; Nelsen, 1999, p. 41]*

A (multivariate) copula is a function C_{M} mapping $[0,1]^m$ to $[0,1]$ satisfying the conditions

(1) $C_{\mathrm{M}}(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for $i = 1, \ldots, m$ and $u_i \in [0,1]$.
(2) $C_{\mathrm{M}}(u_1, \ldots, u_m) = 0$ if $u_i = 0$ for any i.
(3) The range of C_{M} includes zero (grounded).
(4) The volume of C_{M} is non-negative for any m-dimensional interval (m-increasing).

It can also be shown that for any m-dimensional copula greater than or equal to three, each k-dimensional margin of this copula is a k-dimensional copula, for $2 \leq k < m$. □

Hence, a copula is an m-dimensional subcopula whose domain is the entire unit m-dimensional cube $[0,1]^m$. Although, this distinction may appear to be a minor one, it is rather important because every subcopula can be extended to a copula but not in a unique way.

THEOREM 2.11 *[Schweizer and Sklar, 1983, pp. 83–84 (proof); Nelsen, 1999, p. 40]*

Given any subcopula $\widetilde{C}_{\mathrm{M}}$, there is a copula C_{M} such that

$$C_{\mathrm{M}}(u_1, \ldots, u_m) = \widetilde{C}_{\mathrm{M}}(u_1, \ldots, u_m)$$

for all $\mathbf{u}$ in the domain of $\widetilde{C}_{\mathrm{M}}$. □

The range of a subcopula $\widetilde{C}_{\mathrm{M}}$ is also a subset of the unit interval $[0,1]$.

LEMMA 2.12 *[Schweizer and Sklar, 1983, p. 83 (proof); Nelsen, 1999, p. 40]*

An m-dimensional subcopula $\widetilde{C}_{\mathrm{M}}$ is uniformly continuous on its domain and

$$0 \leq \widetilde{C}_{\mathrm{M}}(u_1, \ldots, u_m) \leq \min(u_1, \ldots, u_m)$$

for any $\mathbf{u}$ in the domain of $\widetilde{C}_{\mathrm{M}}$. □

Note that the term "copula" was chosen to emphasize the manner in which a copula "couples" a joint distribution function to its univariate margins.

We are now in a position to relate multivariate distributions to their univariate marginal distributions obtained by integrating over all but one of the dimensions.

THEOREM 2.13 *[Schweizer and Sklar, 1983, pp. 83–84 (proof); Nelsen, 1999, pp. 19 and 41]*

If F_{M} is a multivariate cdf with univariate marginal cdfs $F_1, \ldots, F_m$ then there exists an m-dimensional subcopula $\widetilde{C}_{\mathrm{M}}$, with domain equal to $\mathrm{Ran}(F_1) \times \ldots \times \mathrm{Ran}(F_m)$ where $\mathrm{Ran}(F_i)$ denotes the range of F_i, such that

$$F_{\mathrm{M}}(y_1,\ldots,y_m) = \widetilde{C}_{\mathrm{M}}(F_1(y_1),\ldots,F_m(y_m)) \tag{2.18}$$

for all $y_1,\ldots,y_m \in \mathbf{R}$. The m-dimensional subcopula is given by

$$\widetilde{C}_{\mathrm{M}}(u_1,\ldots,u_m) = F_{\mathrm{M}}(F_1^{-1}(u_1),\ldots,F_m^{-1}(u_m))$$

for all $(u_1,\ldots,u_m)$ in the domain of $\widetilde{C}_{\mathrm{M}}$, where F_j^{-1} is the quantile function corresponding to F_j.
Conversely, let $F_1, \ldots, F_m$ be cdfs, let $\widetilde{C}_{\mathrm{M}}$ be any m-dimensional subcopula whose domain contains the set $\mathrm{Ran}(F_1) \times \ldots \times \mathrm{Ran}(F_m)$ and F_{M} be defined by Equation (2.18). Then F_{M} is a multivariate cdf with margins $F_1, \ldots, F_m$. □

Extension from subcopulas to copulas results as an immediate consequence of Theorems 2.11 and 2.13.

THEOREM 2.14 *[Schweizer and Sklar, 1983, p. 84; Molenberghs, 1992, pp. 61–62; Joe, 1997, pp. 12–13; Nelsen, 1999, pp. 19 and 41]*
Let F_{M} be a multivariate cdf with margins $F_1, \ldots, F_m$. Then there is a copula C_{M} (generally non-unique) such that

$$F_{\mathrm{M}}(y_1,\ldots,y_m) = C_{\mathrm{M}}(F_1(y_1),\ldots,F_m(y_m)) \tag{2.19}$$

for all $y_1,\ldots,y_m \in \mathbf{R}$.
If the functions $F_1, \ldots, F_m$ are all continuous on $\mathbf{R}$ then C_{M} is unique, otherwise C_{M} is uniquely determined on $\mathrm{Ran}(F_1) \times \ldots \times \mathrm{Ran}(F_m)$. □

By means of quantile functions corresponding to the univariate cdfs F_1 to F_m, copulas can be written in terms of uniform univariate margins.

COROLLARY 2.15 *[Schweizer and Sklar, 1983, p. 83; Genest and MacKay, 1986a; Joe, 1997, p. 13; Nelsen, 1999, pp. 19 and 41]*
By letting the univariate margins of the copula be $u_j = F_j(y_j)$ in Equation (2.19), we obtain

$$C_{\mathrm{M}}(u_1,\ldots,u_m) = C_{\mathrm{M}}(\mathbf{u}) = F_{\mathrm{M}}\left(F_1^{-1}(u_1),\ldots,F_m^{-1}(u_m)\right)$$

where the univariate margins $u_1, \ldots, u_m$ are represented by the vector $\mathbf{u}$ and the multivariate cdf F_{M} is now in terms of the quantile functions $F_j^{-1}(u_j) = y_j$. ○

Copulas belong to the Fréchet class $\mathcal{F}$, the class of multivariate distributions with given univariate margins, which allows us to define the bounds.

THEOREM 2.16 *[Joe, 1997, p. 58 (proof); Nelsen, 1999, pp. 8–9, 26, and 42]*
If $F_{\mathrm{M}} \in \mathcal{F}(F_1,\ldots,F_m)$ then for all $\mathbf{y} \in \mathbf{R}^m$,

$$\max\{0, F_1 + \ldots + F_m - (m-1)\} \le F_{\mathrm{M}} \le \min\{F_1,\ldots,F_m\}$$

are the Fréchet lower and upper bounds, respectively F_L and F_U. □

As mentioned earlier, copulas are just dependency structures linking the margins together. Copulas can therefore be considered to be “independent” from their univariate margins (Joe, 1997, p. 13; Nelsen, 1999, p. 21).

THEOREM 2.17 *[Joe, 1997, p. 13; Nelsen, 1999, p. 22 (proof)]*
A copula $C_{\mathrm{M}}(\mathbf{u})$ is invariant under strictly increasing transformations of its random variables $\mathbf{Y}$. □

This theorem can be generalized to strictly monotone transformations which also includes strictly decreasing transformations. In the case where at least one of the transformations is strictly decreasing, the copula is no longer invariant to the transformation but will change in a predictable way yielding an associated copula.

PROPERTY 2.18 *[Joe, 1997, pp. 15–16; Nelsen, 1999, pp. 21–22]*
If a copula has m margins, there are $(2^m - 1)$ ways of applying these strictly decreasing transformations each one yielding an associated copula.
In the bivariate case, the three relationships between a copula and its associated copulas are:

(1) If a transformation is strictly increasing on the range of its random variable Y_1 and a transformation is strictly decreasing on the range of its random variable Y_2, then

$$C_{\mathrm{B}}^{*}(u_1,u_2) = u_1 - C_{\mathrm{B}}(u_1, 1-u_2) = u_1 - C_{\mathrm{B}}(u_1, v_2)$$

(2) By interchanging the transformations applied on the ranges of its random variables Y_1 and Y_2, the second associated copula $C_{\mathrm{B}}^{**}(u_1,u_2)$ is obtained.
(3) If transformations are strictly decreasing on both the range of its random variable Y_1 and the range of its random variable Y_2, then

$$\begin{aligned} C_{\mathrm{B}}^{***}(u_1,u_2) &= u_1 + u_2 - 1 + C_{\mathrm{B}}(1-u_1, 1-u_2) \\ &= u_1 + u_2 - 1 + C_{\mathrm{B}}(v_1, v_2) \end{aligned} \tag{2.20}$$

This can also be extended to the multivariate case using the standard relationship between multivariate cdfs and multivariate survival functions, see Equation (5.35). Note that only a permutation symmetric copula will have m distinct associated copulas. ○

DEFINITION 2.19 *[Joe, 1997, p. 16; Nelsen, 1999, p. 28]*
When a copula is applied to survival functions, the resulting associated copula is called a survival copula. For the bivariate case, this can be seen by rewriting Equation (2.20) as

$$\begin{aligned} \widehat{C}_{\mathrm{B}}(v_1,v_2) &= v_1 + v_2 - 1 + C_{\mathrm{B}}(1-v_1, 1-v_2) \\ &= 1 - u_1 - u_2 + C_{\mathrm{B}}(u_1,u_2) \\ &= \overline{C}_{\mathrm{B}}(u_1,u_2) \end{aligned}$$

where $\widehat{C}_{\mathrm{B}}(v_1,v_2)$ is a bivariate survival copula. Note that the survival copula $\widehat{C}_{\mathrm{M}}(\mathbf{v})$ should not be confused with $\overline{C}_{\mathrm{M}}(\mathbf{u})$, the joint survival function of uniform $U(0,1)$ random variables whose joint cdf is the copula $C_{\mathrm{M}}(\mathbf{u})$. □

A tractable form of a (survival) copula is obtained if it can be written as a sum of functions of the marginals.

$$\psi(C_{\mathrm{M}}(\mathbf{u})) = \psi(u_1) + \ldots + \psi(u_m) = \sum_{j=1}^{m} \psi(u_j)$$

DEFINITION 2.20 *[Frank, 1979; Nelsen, 1999, p. 90]*
If ψ is a continuous strictly decreasing function mapping $[0,1]$ to $[0,\infty]$ such that $\psi(1) = 0$, then its pseudo-inverse is the function $\psi^{[-1]}$ which maps $[0,\infty]$ to $[0,1]$. It is defined as follows.

$$\psi^{[-1]}(x) = \begin{cases} \psi^{-1}(x) & 0 \leq x \leq \psi(0) \\ 0 & \psi(0) \leq x \leq \infty \end{cases}$$

Hence, it is continuous and non-increasing on $[0,\infty]$, and strictly decreasing on $[0,\psi(0)]$. Furthermore, $\psi^{[-1]}(\psi(x)) = x$ on $[0,1]$, and

$$\begin{aligned} \psi(\psi^{[-1]}(x)) &= \begin{cases} x & 0 \leq x \leq \psi(0) \\ \psi(0) & \psi(0) \leq x \leq \infty \end{cases} \\ &= \min(x, \psi(0)) \end{aligned}$$

Finally, if $\psi(0) = \infty$ then $\psi^{[-1]}(x) = \psi^{-1}(x)$, the inverse function. □

Additional constraints are necessary to ensure that a multivariate distribution is indeed obtained and is also a copula.

LEMMA 2.21 *[Frank, 1979; Nelsen, 1999, pp. 90–91 (proof)]*
Let ψ be a continuous strictly decreasing function mapping $[0,1]$ to $[0,\infty]$ such that $\psi(1) = 0$, and let $\psi^{[-1]}$ be the pseudo-inverse of ψ as just defined. If $G_{\mathrm{M}}(\mathbf{x})$ is a function mapping $[0,1]^m$ to $[0,1]$ given by

$$G_{\mathrm{M}}(\mathbf{x}) = \psi^{[-1]}\left(\sum_{j=1}^{m} \psi(x_j)\right)$$

then $G_{\mathrm{M}}(\mathbf{x})$ satisfies the boundary conditions for a copula. But this multivariate function is a copula if and only if ψ is convex. □

This provides a simple class of copulas which has been widely used in the literature.

DEFINITION 2.22 *[Joe, 1997, p. 87; Nelsen, 1999, pp. 92 and 98]*
A (survival) copula of the form

$$C_{\mathrm{M}}(\mathbf{u}) = \psi^{[-1]}\left(\sum_{j=1}^{m} \psi(u_j)\right) \tag{2.21}$$

is called an Archimedean (survival) copula. Notice that it is permutation symmetric in its m margins. □

The meaning of the term Archimedean comes from the similarities with the Archimedean axiom for the positive real numbers.

LEMMA 2.23 *[Frank, 1979 (proof)]*

Similarly, there exists a continuous strictly increasing function mapping $[0,1]$ to $[0,\infty]$ such that $\xi(0)=0$ with pseudo-inverse $\xi^{[-1]}$ given by

$$\xi^{[-1]}(x) = \begin{cases} \xi^{-1}(x) & 0 \leq x \leq \xi(1) \\ 0 & \xi(1) \leq x \leq \infty \end{cases}$$

If $\xi(1)=\infty$ then $\xi^{[-1]}(x)=\xi^{-1}(x)$.

If $G_{\mathrm{M}}(\mathbf{x})$ is a function mapping $[0,1]^m$ to $[0,1]$ given by

$$G_{\mathrm{M}}(\mathbf{x}) = \xi^{[-1]}\left(\sum_{j=1}^{m}\xi(x_j)\right)$$

then $G_{\mathrm{M}}(\mathbf{x})$ satisfies the boundary conditions for a copula. But this multivariate function is a copula if and only if ξ is convex.

Finally, a (survival) copula of the form

$$C_{\mathrm{M}}(\mathbf{u}) = \xi^{[-1]}\left(\sum_{j=1}^{m}\xi(u_j)\right)$$

is also called an Archimedean (survival) copula. □

DEFINITION 2.24 *[Frank, 1979; Nelsen, 1999, p. 92]*

Note that the functions ψ and ξ are called additive generators of the copula. For instance, if we set $\gamma(x) = \exp(-\psi(x))$ and $\gamma^{[-1]}(x) = \psi^{[-1]}(-\ln(x))$ we then obtain

$$C_{\mathrm{M}}(\mathbf{u}) = \gamma^{[-1]}\left(\prod_{j=1}^{m}\gamma(u_j)\right)$$

so that γ is a multiplicative generator. □

If $\psi(0)=\infty$ then ψ is called a strict generator. In this case, $\psi^{[-1]}=\psi^{-1}$ and $C_{\mathrm{M}}(\mathbf{u}) = \psi^{-1}\left(\sum_{j=1}^{m}\psi(u_j)\right)$ is said to be a strict Archimedean (survival) copula.

A subclass of Archimedean (survival) copulas is obtained when the additive generator ϕ is chosen to be a univariate inverse Laplace transform. Recall from Property 2.3 that they are decreasing convex functions (Joe, 1997, p. 373; Nelsen, 1999, p. 106).

Thus, a multivariate cdf or Archimedean copula

$$F_{\mathrm{M}}(\mathbf{u}) = C_{\mathrm{M}}(\mathbf{u}) = \phi\left(\sum_{j=1}^{m}\phi^{-1}(u_j)\right) \tag{2.22}$$

can directly be obtained from Equation (2.21). Equivalently, a multivariate survival function or Archimedean survival copula

$$S_{\mathrm{M}}(\mathbf{v}) = \widehat{C}_{\mathrm{M}}(\mathbf{v}) = \phi\left(\sum_{j=1}^{m} \phi^{-1}(v_j)\right) \tag{2.23}$$

can be obtained.

The methods described in this subsection create general families of multivariate distributions where the multivariate structure can be represented by a (survival) copula. Because the dependence structure is invariant to transformations of the margins, care must taken not to confuse different multivariate distributions.

Indeed, it is important to know exactly what margins have been inserted into a (survival) copula. The simplest example is to consider a survival copula $\widehat{C}_{\mathrm{M}}(\mathbf{y})$ where the margins are survival functions and its associated copula $C_{\mathrm{M}}(\mathbf{y})$ where the margins are cdfs. These are respectively a multivariate survival function and a multivariate cdf. Although both have identical dependence structures, the corresponding densities obtained by taking partial derivatives with respect to each of the margins will generally be different.

2.2.3 MULTIVARIATE LAPLACE TRANSFORMS

The last method to be considered consists in extending the definition of a Laplace transform to the multivariate case.

DEFINITION 2.25 *[Abramowitz and Stegun, 1965, p. 1020; Hougaard, 2000, p. 498]* A bivariate Laplace transform is

$$\phi(s_1, s_2) = \int_0^\infty \mathrm{e}^{-s_1\lambda}\mathrm{e}^{-s_2\lambda}\mathrm{d}F(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda(s_1+s_2)}\mathrm{d}F(\lambda) = \phi(s_1 + s_2)$$

where the two univariate exponential survival functions have λ as common parameter and s_j is either a (transformed) cdf or survival function. Note that bivariate Laplace transforms can in some cases be interpreted as bivariate cdfs or survival functions. □

This relationship can also be generalized to m-dimensions,

$$\phi(\mathbf{s}) = \int_0^\infty \mathrm{e}^{-\lambda\left(\sum_{j=1}^{m} s_j\right)}\mathrm{d}F(\lambda) = \phi\left(\sum_{j=1}^{m} s_j\right) \tag{2.24}$$

where $\phi(\mathbf{s})$ is a multivariate Laplace transform.

Depending on the choice of s_j, these relationships can generate interesting multivariate distributions.

For instance by choosing s_j to be $-\ln[v_j]$, minus the logarithm of an arbitrary survival function corresponding to the integrated intensity $\breve{v}_j = H_j(y_j)$, yields a multivariate survival distribution (Hougaard, 2000, p. 222).

$$S_{\mathrm{M}}(\breve{\mathbf{v}}) = \phi\left(-\sum_{j=1}^{m} \ln[v_j]\right) = \phi\left(\sum_{j=1}^{m} \breve{v}_j\right) \tag{2.25}$$

On the other hand, choosing s_j to be $-\ln[u_j]$, minus the logarithm of an arbitrary cdf, yields a multivariate cdf.

$$F_{\mathrm{M}}(-\ln[u_1],\ldots,-\ln[u_m]) = \phi\left(-\sum_{j=1}^{m}\ln[u_j]\right) \tag{2.26}$$

It can also be seen that Equations (2.25) and (2.26) can respectively be rewritten as a mixture of powers of survival functions and cdfs as was done in Equations (2.12) and (2.14). Now respectively inserting Equations (2.13) and (2.15) yields

$$\begin{aligned} S_{\mathrm{M}}(\mathbf{v}) &= \int_0^\infty \prod_{j=1}^{m} \widetilde{v}_j^{\lambda} \mathrm{d}F(\lambda) = \phi\left(-\sum_{j=1}^{m}\ln[\tilde{v}_j]\right) = \phi\left(\sum_{j=1}^{m}\phi^{-1}(v_j)\right) \\ F_{\mathrm{M}}(\mathbf{u}) &= \int_0^\infty \prod_{j=1}^{m} \widetilde{u}_j^{\lambda} \mathrm{d}F(\lambda) = \phi\left(-\sum_{j=1}^{m}\ln[\tilde{u}_j]\right) = \phi\left(\sum_{j=1}^{m}\phi^{-1}(u_j)\right) \end{aligned} \tag{2.27}$$

which are respectively Equations (2.23) and (2.22). Thus, this can be seen as another method to obtain the subclass of Archimedean (survival) copulas mentioned earlier (Joe, 1997, pp. 86–87; Nelsen, 1999, pp. 65–66).

Finally, notice that Equations (2.25) and (2.26) could respectively be obtained by inserting mixture margins of the form of Equation (2.12) into an Archimedean survival copula and of the form of Equation (2.14) into an Archimedean copula.

2.3 Further reading

3
Closed form mixtures

This chapter illustrates certain possible examples of closed form mixtures that will be used to induce dependencies in Chapter 5. Many of these are well known because they have been used widely to handle over-dispersion for univariate data. The two first sections cover mixture cdfs and mixture survival functions, whereas the third section contains mixture densities. Each of the subsections correspond to the mixing distribution used to provide certain properties which will later characterize the type of dependence induced.

Due to the close relationship between cdfs and survival functions presented in Equation (2.6), specific mixtures are only presented in one of the two first sections. Although the mixture density corresponding to a specific mixture cdf and survival function (or vice-versa) can be obtained from the relationships given in Appendix B.1, each of these is not automatically the result of a mixture. Hence, the two first sections are somewhat complementary whereas the third only has the corresponding mixtures for certain particular combinations of distributions but includes additional cases.

Many of the mixtures considered can be related to Laplace transforms. Although the classical method consists in choosing the conditional survival function of Equation (2.4) to be an exponential distribution, this must not necessarily be the case. Indeed, any mixture involving a conditional or mixing distribution which includes an exponential of y can be regarded as a Laplace transform.

Certain mathematical functions given in Section B.2 are required in order to specify the different distributions and mixtures. These are $\Gamma(\cdot)$ the gamma function, $K(\cdot,\cdot)$ the fractional Bessel function of the third kind, and $B(\cdot)$ the beta function respectively defined in Equations (B.1), (B.2), and (B.3).

Finally, each closed form mixture is obtained by rearranging the terms of the integral into two distinct parts. The first part includes all the terms involving the variable of integration, always defined as λ in the below developments. Additional constants with respect to the variable of integration are then introduced in order to obtain a known distribution. The second part consists of the remaining terms which do not involve the variable of integration and can therefore come out of the integral. Hence, the second part corresponds to the resulting mixture as the remaining integral is a distribution and evaluates to unity. The distribution integrated out is each time mentioned as it is of major interest in Chapter 5 along with the notions of "closed under sampling" and "recursively closed" (respectively given in Definitions 5.2 and 5.5).

3.1 Cumulative distribution functions

The distributions obtained in this section are mixture cdfs resulting from Equation (2.2).

3.1.1 EXPONENTIAL

The following mixture has a one parameter exponential density, Equation (B.5), as mixing distribution. The closed form mixture is obtained by integrating out a re-parameterized one parameter exponential distribution (recursively closed).

Gumbel
The conditional cdf is a two parameter Gumbel, Equation (B.16), where the location is re-parameterized as $\alpha+\beta\ln[\lambda\beta]$ and the scale is set equal to β.

$$\begin{aligned}
F_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\phi\mathrm{e}^{-\phi\lambda}\mathrm{d}\lambda \\
&= \int_0^\infty \phi\mathrm{e}^{-\lambda\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\phi}{\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)\mathrm{e}^{-\lambda\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}\mathrm{d}\lambda \\
&= \frac{\phi}{\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}} \\
&= \frac{\phi\mathrm{e}^{\frac{y-\alpha}{\beta}}}{\beta+\phi\mathrm{e}^{\frac{y-\alpha}{\beta}}}
\end{aligned}$$

When $\phi=\beta$, the resulting mixture has the form of a two parameter logistic cdf, Equation (B.27).
(Johnson and Kotz, 1970b, p. 3)

3.1.2 GAMMA

The following mixture has a two parameter gamma density, Equation (B.8), as mixing distribution. The closed form mixture is obtained by integrating out a re-parameterized two parameter gamma distribution (recursively closed).

Gumbel
The conditional cdf is a two parameter Gumbel, Equation (B.16), where the location is re-parameterized as $\psi+\phi\ln[\lambda\phi]$ and the scale is set equal to ϕ.

$$F_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda\phi\mathrm{e}^{-\frac{y-\psi}{\phi}}}\frac{\beta^\alpha\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}}{\Gamma(\alpha)}\mathrm{d}\lambda$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\left(\beta+\phi \mathrm{e}^{-\frac{y-\psi}{\phi}}\right)}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha}{\left(\beta+\phi \mathrm{e}^{-\frac{y-\psi}{\phi}}\right)^\alpha} \frac{\left(\beta+\phi \mathrm{e}^{-\frac{y-\psi}{\phi}}\right)^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\left(\beta+\phi \mathrm{e}^{-\frac{y-\psi}{\phi}}\right)}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \left(\frac{\beta}{\beta+\phi \mathrm{e}^{-\frac{y-\psi}{\phi}}}\right)^\alpha
\end{aligned}
$$

The resulting mixture has the form of a four parameter generalized logistic cdf, Equation (B.31).
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

3.1.3 INVERSE GAUSSIAN

The following mixture has a two parameter inverse Gaussian density, Equation (B.9), as mixing distribution. The closed form mixture is obtained by integrating out a re-parameterized two parameter inverse Gaussian distribution (recursively closed for one parameter inverse Gaussian distribution where $\phi = \psi$).

Gumbel

The conditional cdf is a two parameter Gumbel, Equation (B.16), where the location is re-parameterized as $\alpha + \beta \ln[\lambda\beta]$ and the scale is set equal to β.

$$
\begin{aligned}
F_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}\left(\phi+2\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{a}{2\pi\gamma^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi\left(\phi+2\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}}
\end{aligned}
$$

Where $a = \sqrt{\psi\left(\phi + 2\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}$ and $\gamma = \frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda = \frac{\psi}{a}\mathrm{d}\gamma$.

3.1.4 GENERALIZED INVERSE GAUSSIAN

The following mixture has a three parameter generalized inverse Gaussian density, Equation (B.10), as mixing distribution. The closed form mixture is obtained by integrating out a re-parameterized three parameter generalized inverse Gaussian distribution (recursively closed).

Gumbel
The conditional cdf is a two parameter Gumbel, Equation (B.16), where the location is re-parameterized as $\alpha + \beta \ln[\lambda\beta]$ and the scale is set equal to β.

$$
\begin{aligned}
F_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\left(\frac{\psi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}, |\epsilon|\right)}{\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \\
&\qquad \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}, |\epsilon|\right)} \mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \left(\frac{\phi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\right)^{\frac{\epsilon}{2}}
\end{aligned}
$$

3.2 Survival functions

The distributions obtained in this section are mixture survival functions resulting from Equation (2.3). The special case of Laplace transform occurs when the conditional survival function is chosen to be an exponential distribution, Equation (B.6).

A generalization of survival mixtures is obtained if the argument y of the conditional Weibull survival function is replaced as by $G(y)$, an arbitrary function of

y. Note that this function $G(y)$ must therefore be a strictly increasing transformation of y in order not to affect the range of the survival mixture. The conditional survival functions will then be of form

$$\begin{gathered} \mathrm{e}^{-\lambda G(y)^\kappa} \\ \mathrm{e}^{-[\lambda G(y)]^\kappa} \end{gathered} \tag{3.1}$$

where $G(y)$ defines the distribution. For instance, the Weibull, Gompertz, and generalized Gompertz distributions are obtained when $G(y)$ is respectively chosen to be y, e^y, and e^{y^σ}.

3.2.1 EXPONENTIAL

The following mixtures are all based on one parameter exponential densities, Equation (B.5), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized one parameter exponential distribution (recursively closed).

Exponential

The conditional survival function is a one parameter exponential, Equation (B.6), resulting in an exponential Laplace transform.

$$\begin{aligned} S_\mathrm{m}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y}\beta\mathrm{e}^{-\lambda\beta}\mathrm{d}\lambda \\ &= \int_0^\infty \beta\mathrm{e}^{-\lambda(\beta+y)}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta}{\beta+y}(\beta+y)\mathrm{e}^{-\lambda(\beta+y)}\mathrm{d}\lambda \\ &= \frac{\beta}{\beta+y} \end{aligned}$$

Weibull

The conditional survival function is a two parameter Weibull, Equation (B.13), where θ^α has been re-parameterized to λ.

$$\begin{aligned} S_\mathrm{m}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y^\kappa}\beta\mathrm{e}^{-\lambda\beta}\mathrm{d}\lambda \\ &= \int_0^\infty \beta\mathrm{e}^{-\lambda[\beta+y^\kappa]}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta}{\beta+y^\kappa}[\beta+y^\kappa]\mathrm{e}^{-\lambda[\beta+y^\kappa]}\mathrm{d}\lambda \\ &= \frac{\beta}{\beta+y^\kappa} \end{aligned}$$

Gompertz

The conditional survival function is a two parameter Gompertz, Equation (B.18), where θ^{ϵ} has been re-parameterized to λ.

$$\begin{aligned} S_{\mathrm{m}}(y) &= \int_0^{\infty} \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y}} \beta \mathrm{e}^{-\lambda\beta} \mathrm{d}\lambda \\ &= \int_0^{\infty} \beta \mathrm{e}^{-\lambda(\beta+\mathrm{e}^{\kappa y})} \mathrm{d}\lambda \\ &= \int_0^{\infty} \frac{\beta}{\beta+\mathrm{e}^{\kappa y}} \left(\beta+\mathrm{e}^{\kappa y}\right) \mathrm{e}^{-\lambda(\beta+\mathrm{e}^{\kappa y})} \mathrm{d}\lambda \\ &= \frac{\beta}{\beta+\mathrm{e}^{\kappa y}} \end{aligned}$$

When $\beta = 1$, the resulting mixture has the form of a one parameter logistic cdf, Equation (B.28). In such a case, this also corresponds to the mixture cdf obtained when a Gumbel density is used.

Generalized Gompertz

The conditional survival function is a three parameter generalized Gompertz, Equation (B.22), where θ^{ϵ} has been re-parameterized to λ.

$$\begin{aligned} S_{\mathrm{m}}(y) &= \int_0^{\infty} \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y^{\nu}}} \beta \mathrm{e}^{-\lambda\beta} \mathrm{d}\lambda \\ &= \int_0^{\infty} \beta \mathrm{e}^{-\lambda\left(\beta+\mathrm{e}^{\kappa y^{\nu}}\right)} \mathrm{d}\lambda \\ &= \int_0^{\infty} \frac{\beta}{\beta+\mathrm{e}^{\kappa y^{\nu}}} \left(\beta+\mathrm{e}^{\kappa y^{\nu}}\right) \mathrm{e}^{-\lambda\left(\beta+\mathrm{e}^{\kappa y^{\nu}}\right)} \mathrm{d}\lambda \\ &= \frac{\beta}{\beta+\mathrm{e}^{\kappa y^{\nu}}} \end{aligned}$$

Generalization

The conditional survival function is chosen to have the form presented in Equation (3.1).

$$\begin{aligned} S_{\mathrm{m}}(y) &= \int_0^{\infty} \mathrm{e}^{-\lambda G(y)^{\kappa}} \beta \mathrm{e}^{-\lambda\beta} \mathrm{d}\lambda \\ &= \int_0^{\infty} \beta \mathrm{e}^{-\lambda[\beta+G(y)^{\kappa}]} \mathrm{d}\lambda \\ &= \int_0^{\infty} \frac{\beta}{\beta+G(y)^{\kappa}} \left[\beta+G(y)^{\kappa}\right] \mathrm{e}^{-\lambda[\beta+G(y)^{\kappa}]} \mathrm{d}\lambda \\ &= \frac{\beta}{\beta+G(y)^{\kappa}} \end{aligned}$$

3.2.2 GAMMA

The following mixtures are all based on two parameter gamma densities, Equation (B.8), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized two parameter gamma distribution (recursively closed).

Exponential
The conditional survival function is a one parameter exponential, Equation (B.6), resulting in a gamma Laplace transform.

$$\begin{aligned} S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+y]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha}{[\beta+y]^\alpha} \frac{[\beta+y]^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+y]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \left[\frac{\beta}{\beta+y}\right]^\alpha \end{aligned}$$

The resulting mixture has the form of a two parameter Pareto survival function, Equation (B.24).
(Johnson and Kotz, 1970a, pp. 233–234; Cox and Oakes, 1984, pp. 19–20)

Weibull
The conditional survival function is a two parameter Weibull, Equation (B.13), where θ^α has been re-parameterized to λ.

$$\begin{aligned} S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y^\kappa} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+y^\kappa]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha}{[\beta+y^\kappa]^\alpha} \frac{[\beta+y^\kappa]^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+y^\kappa]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \left[\frac{\beta}{\beta+y^\kappa}\right]^\alpha \end{aligned}$$

This mixture has the form of a three parameter Burr distribution, Equation (B.37).
(Johnson and Kotz, 1970a, p. 266)

Gompertz
The conditional survival function is a two parameter Gompertz, Equation (B.18), where θ^ϵ has been re-parameterized to λ.

$$S_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y}} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+\mathrm{e}^{\kappa y}]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha}{[\beta+\mathrm{e}^{\kappa y}]^\alpha} \frac{[\beta+\mathrm{e}^{\kappa y}]^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+\mathrm{e}^{\kappa y}]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \left[\frac{\beta}{\beta+\mathrm{e}^{\kappa y}}\right]^\alpha
\end{aligned}
$$

Generalized Gompertz

The conditional survival function is a three parameter generalized Gompertz, Equation (B.22), where θ^ϵ has been re-parameterized to λ.

$$
\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y^\nu}} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\left[\beta+\mathrm{e}^{\kappa y^\nu}\right]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha}{[\beta+\mathrm{e}^{\kappa y^\nu}]^\alpha} \frac{\left[\beta+\mathrm{e}^{\kappa y^\nu}\right]^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\left[\beta+\mathrm{e}^{\kappa y^\nu}\right]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \left[\frac{\beta}{\beta+\mathrm{e}^{\kappa y^\nu}}\right]^\alpha
\end{aligned}
$$

Generalization

The conditional survival function is chosen to have the form presented in Equation (3.1).

$$
\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda G(y)^\kappa} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha}{[\beta+G(y)^\kappa]^\alpha} \frac{[\beta+G(y)^\kappa]^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \left[\frac{\beta}{\beta+G(y)^\kappa}\right]^\alpha
\end{aligned}
$$

3.2.3 INVERSE GAUSSIAN

The following mixtures are all based on two parameter inverse Gaussian densities, Equation (B.9), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized one parameter inverse Gaussian distribution (recursively closed for one parameter inverse Gaussian distribution where $\phi = \psi$).

Exponential
The conditional survival function is a one parameter exponential, Equation (B.6), resulting in a inverse Gaussian Laplace transform.

$$
\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\frac{\psi}{a}\mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{a}{2\pi\gamma^3}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi(\phi+2y)}}
\end{aligned}
$$

Where $a=\sqrt{\psi(\phi+2y)}$ and $\gamma=\frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda=\frac{\psi}{a}\mathrm{d}\gamma$.

Weibull
The conditional survival function is a two parameter Weibull, Equation (B.13), where θ^α has been re-parameterized to λ.

$$
\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y^\kappa}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}(\phi+2y^\kappa)-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\frac{\psi}{a}\mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{a}{2\pi\gamma^3}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi(\phi+2y^\kappa)}}
\end{aligned}
$$

Where $a=\sqrt{\psi\left(\phi+2y^\kappa\right)}$, $\gamma=\frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda=\frac{\psi}{a}\mathrm{d}\gamma$.

Gompertz

The conditional survival function is a two parameter Gompertz, Equation (B.18), where θ^{ϵ} has been re-parameterized to λ.

$$\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y}} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}(\phi+2\mathrm{e}^{\kappa y})-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{a}{2\pi\gamma^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi(\phi+2\mathrm{e}^{\kappa y})}}
\end{aligned}$$

Where $a = \sqrt{\psi\,(\phi + 2\mathrm{e}^{\kappa y})}$, $\gamma = \frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda = \frac{\psi}{a}\mathrm{d}\gamma$.

Generalized Gompertz

The conditional survival function is a three parameter generalized Gompertz, Equation (B.22), where θ^{ϵ} has been re-parameterized to λ.

$$\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y^{\nu}}} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}\left(\phi+2\mathrm{e}^{\kappa y^{\nu}}\right)-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{a}{2\pi\gamma^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a}
\end{aligned}$$

$$= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa}y^{\nu}\right)}}$$

Where $a = \sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa}y^{\nu}\right)}$, $\gamma = \frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda = \frac{\psi}{a}\mathrm{d}\gamma$.

Generalization
The conditional survival function is chosen to have the form presented in Equation (3.1).

$$\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda G(y)^{\kappa}} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}[\phi+2G(y)^{\kappa}]-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{a}{2\pi\gamma^3}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \mathrm{d}\gamma \\
&= \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi[\phi+2G(y)^{\kappa}]}}
\end{aligned}$$

Where $a = \sqrt{\psi\left[\phi+2G(y)^{\kappa}\right]}$, and $\gamma = \frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda = \frac{\psi}{a}\mathrm{d}\gamma$.

3.2.4 GENERALIZED INVERSE GAUSSIAN

The following mixtures are all based on three parameter generalized inverse Gaussian densities, Equation (B.10), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized three parameter generalized inverse Gaussian distribution (recursively closed).

Exponential
The conditional survival function is a one parameter exponential, Equation (B.6), resulting in a generalized inverse Gaussian Laplace transform.

$$\begin{aligned}
S_{\mathrm{m}}(y) &= \int_0^\infty \mathrm{e}^{-\lambda y} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda
\end{aligned}$$

$$= \int_0^\infty \frac{K\left(\sqrt{\psi(\phi+2y)}, |\epsilon|\right)\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon}{2}}}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi(\phi+2y)}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \frac{K\left(\sqrt{\psi(\phi+2y)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \left(\frac{\phi}{\phi+2y}\right)^{\frac{\epsilon}{2}}$$

Weibull

The conditional survival function is a two parameter Weibull, Equation (B.13), where θ^α has been re-parameterized to λ.

$$S_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda y^\kappa} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2y^\kappa)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{K\left(\sqrt{\psi(\phi+2y^\kappa)}, |\epsilon|\right)\left(\frac{\psi}{\phi+2y^\kappa}\right)^{\frac{\epsilon}{2}}}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2y^\kappa)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2y^\kappa}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi(\phi+2y^\kappa)}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \frac{K\left(\sqrt{\psi(\phi+2y^\kappa)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \left(\frac{\phi}{\phi+2y^\kappa}\right)^{\frac{\epsilon}{2}}$$

Gompertz

The conditional survival function is a two parameter Gompertz, Equation (B.18), where θ^ϵ has been re-parameterized to λ.

$$S_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y}} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2\mathrm{e}^{\kappa y})-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y}\right)}, |\epsilon|\right)\left(\frac{\psi}{\phi+2\mathrm{e}^{\kappa y}}\right)^{\frac{\epsilon}{2}}}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\mathrm{e}^{\kappa y}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2\mathrm{e}^{\kappa y}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y}\right)}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \frac{K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y}\right)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)}\left(\frac{\phi}{\phi+2\mathrm{e}^{\kappa y}}\right)^{\frac{\epsilon}{2}}$$

Generalized Gompertz

The conditional survival function is a three parameter generalized Gompertz, Equation (B.22), where θ^ϵ has been re-parameterized to λ.

$$S_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda \mathrm{e}^{\kappa y^\nu}} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\mathrm{e}^{\kappa y^\nu}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y^\nu}\right)}, |\epsilon|\right)\left(\frac{\psi}{\phi+2\mathrm{e}^{\kappa y^\nu}}\right)^{\frac{\epsilon}{2}}}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\mathrm{e}^{\kappa y^\nu}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2\mathrm{e}^{\kappa y^\nu}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y^\nu}\right)}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \frac{K\left(\sqrt{\psi\left(\phi+2\mathrm{e}^{\kappa y^\nu}\right)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)}\left(\frac{\phi}{\phi+2\mathrm{e}^{\kappa y^\nu}}\right)^{\frac{\epsilon}{2}}$$

Generalization

The conditional survival function is chosen to have the form presented in Equation (3.1).

$$S_{\mathrm{m}}(y) = \int_0^\infty \mathrm{e}^{-\lambda G(y)^\kappa} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2G(y)^\kappa\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \int_0^\infty \frac{K\left(\sqrt{\psi\left(\phi+2G(y)^\kappa\right)}, |\epsilon|\right)\left(\frac{\psi}{\phi+2G(y)^\kappa}\right)^{\frac{\epsilon}{2}}}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2G(y)^\kappa\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2G(y)^\kappa}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\left(\phi+2G(y)^\kappa\right)}, |\epsilon|\right)} \mathrm{d}\lambda$$

$$= \frac{K\left(\sqrt{\psi\left(\phi+2G(y)^\kappa\right)}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)}\left(\frac{\phi}{\phi+2G(y)^\kappa}\right)^{\frac{\epsilon}{2}}$$

3.3 Densities

The distributions obtained in this section are mixture densities resulting from Equation (2.1). A generalization is obtained if the argument y of the conditional Weibull density is replaced as by $G(y)$, an arbitrary function of y. Note that this function $G(y)$ must therefore be a strictly increasing transformation of y in order not to affect the range of the mixture density. The conditional densities will then be of form

$$\begin{gathered} \kappa\lambda g(y)G(y)^{\kappa-1}\mathrm{e}^{-\lambda G(y)^\kappa} \\ \kappa\lambda^\kappa g(y)G(y)^{\kappa-1}\mathrm{e}^{-\lambda^\kappa G(y)^\kappa} \end{gathered} \tag{3.2}$$

where $G(y)$ defines the distribution and $\frac{\partial G(y)}{\partial y} = g(y)$ is the corresponding Jacobian. For instance, the Weibull, Gompertz, and generalized Gompertz distributions are obtained when $G(y)$ is respectively chosen to be y, e^y, and e^{y^σ}.

3.3.1 GAUSSIAN

The following mixtures are all based on two parameter Gaussian densities, Equation (B.41), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized two parameter Gaussian distribution (closed under sampling).

Gaussian
The conditional density is a two parameter Gaussian, Equation (B.41).

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\frac{(y-\lambda)^2}{2\psi^2}}}{\sqrt{2\pi\psi^2}} \frac{\mathrm{e}^{-\frac{(\lambda-\kappa)^2}{2\nu^2}}}{\sqrt{2\pi\nu^2}} \mathrm{d}\lambda \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\frac{y^2}{2\psi^2}+\frac{y\lambda}{\psi^2}-\frac{\lambda^2}{2\psi^2}}}{\sqrt{2\pi\psi^2}} \frac{\mathrm{e}^{-\frac{\kappa^2}{2\nu^2}+\frac{\lambda\kappa}{\nu^2}-\frac{\lambda^2}{2\nu^2}}}{\sqrt{2\pi\nu^2}} \mathrm{d}\lambda \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\lambda^2\frac{\nu^2+\psi^2}{2\nu^2\psi^2}+\lambda\frac{\psi^2\kappa+\nu^2 y}{\nu^2\psi^2}-\frac{\kappa^2\psi^2+\nu^2 y^2}{2\nu^2\psi^2}}}{2\pi\sqrt{\psi^2\nu^2}} \mathrm{d}\lambda \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\frac{\nu^2+\psi^2}{2\nu^2\psi^2}\left[\lambda^2 - 2\lambda\frac{\psi^2\kappa+\nu^2 y}{\nu^2+\psi^2} + \frac{(\psi^2\kappa+\nu^2 y)^2}{(\nu^2+\psi^2)^2} - \frac{(\psi^2\kappa+\nu^2 y)^2}{(\nu^2+\psi^2)^2} + \frac{\kappa^2\psi^2+\nu^2 y^2}{\nu^2+\psi^2}\right]}}{\sqrt{2\pi}\sqrt{2\pi\psi^2\nu^2}}\mathrm{d}\lambda \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\frac{\nu^2+\psi^2}{2\nu^2\psi^2}\left[\frac{\kappa^2\psi^2+\nu^2 y^2}{\nu^2+\psi^2} - \frac{(\psi^2\kappa+\nu^2 y)^2}{(\nu^2+\psi^2)^2}\right]}}{\sqrt{2\pi(\nu^2+\psi^2)}} \\
&\qquad \frac{\sqrt{\nu^2+\psi^2}\mathrm{e}^{-\frac{\nu^2+\psi^2}{2\nu^2\psi^2}\left[\lambda - \frac{\psi^2\kappa+\nu^2 y}{\nu^2+\psi^2}\right]^2}}{\sqrt{2\pi\psi^2\nu^2}}\mathrm{d}\lambda \\
&= \frac{\mathrm{e}^{-\frac{\kappa^2\psi^2+\nu^2 y^2}{2\nu^2\psi^2} + \frac{(\psi^2\kappa+\nu^2 y)^2}{2\nu^2\psi^2(\nu^2+\psi^2)}}}{\sqrt{2\pi(\nu^2+\psi^2)}} \\
&= \frac{\mathrm{e}^{-\frac{y^2}{2(\nu^2+\psi^2)} + \frac{y\kappa}{(\nu^2+\psi^2)} - \frac{\kappa^2}{2(\nu^2+\psi^2)}}}{\sqrt{2\pi(\nu^2+\psi^2)}} \\
&= \frac{\mathrm{e}^{-\frac{(y-\kappa)^2}{2(\nu^2+\psi^2)}}}{\sqrt{2\pi(\nu^2+\psi^2)}}
\end{aligned}
$$

The resulting mixture has the form of a two parameter Gaussian density, Equation (B.41).
(Johnson and Kotz, 1970a, pp. 87–88)

3.3.2 EXPONENTIAL

The following mixtures are all based on one parameter exponential densities, Equation (B.5), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized two parameter gamma distribution, Equation (B.8). Note that in the case of an exponential, Gumbel, or general type of density presented in Equation (3.2), the shape parameter α is fixed at two.

Poisson

The conditional density is a Poisson, Equation (B.45).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^{\infty} \frac{\lambda^y \mathrm{e}^{-\lambda}}{y!}\beta\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda \\
&= \int_0^{\infty} \frac{\beta\lambda^y\mathrm{e}^{-\lambda(\beta+1)}}{\Gamma(y+1)}\mathrm{d}\lambda \\
&= \int_0^{\infty} \frac{\beta}{(\beta+1)^{y+1}}\frac{(\beta+1)^{y+1}\lambda^y\mathrm{e}^{-\lambda(\beta+1)}}{\Gamma(y+1)}\mathrm{d}\lambda \\
&= \frac{\beta}{(\beta+1)^{y+1}}
\end{aligned}
$$

Gamma

The conditional density is a two parameter gamma, Equation (B.8).

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^\alpha y^{\alpha-1}\mathrm{e}^{-\lambda y}}{\Gamma(\alpha)}\beta\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta\lambda^\alpha y^{\alpha-1}\mathrm{e}^{-\lambda(\beta+y)}}{\Gamma(\alpha)}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\alpha\beta y^{\alpha-1}}{(\beta+y)^{\alpha+1}}\frac{(\beta+y)^{\alpha+1}\lambda^\alpha\mathrm{e}^{-\lambda(\beta+y)}}{\Gamma(\alpha+1)}\mathrm{d}\lambda \\ &= \frac{\alpha\beta y^{\alpha-1}}{(\beta+y)^{\alpha+1}} \end{aligned}$$

Exponential

The conditional density is a one parameter exponential, Equation (B.5), resulting in the mixture density corresponding to an exponential Laplace transform.

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \lambda\mathrm{e}^{-\lambda y}\beta\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda \\ &= \int_0^\infty \beta\lambda\mathrm{e}^{-\lambda(\beta+y)}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta}{(\beta+y)^2}(\beta+y)^2\lambda\mathrm{e}^{-\lambda(\beta+y)}\mathrm{d}\lambda \\ &= \frac{\beta}{(\beta+y)^2} \end{aligned}$$

Generalization

The conditional density is chosen to have the form presented in Equation (3.2).

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \kappa\lambda g(y)G(y)^{\kappa-1}\mathrm{e}^{-\lambda G(y)^\kappa}\beta\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda \\ &= \int_0^\infty \beta\kappa g(y)G(y)^{\kappa-1}\lambda\mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}\mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta\kappa g(y)G(y)^{\kappa-1}}{[\beta+G(y)^\kappa]^2}\left[\beta+G(y)^\kappa\right]^2\lambda\mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}\mathrm{d}\lambda \\ &= \frac{\beta\kappa g(y)G(y)^{\kappa-1}}{[\beta+G(y)^\kappa]^2} \end{aligned}$$

Gumbel

The conditional density is a two parameter Gumbel, Equation (B.15), where the location is re-parameterized as $\alpha+\beta\ln[\lambda\beta]$ and the scale is set equal to β.

$$f_{\mathrm{m}}(y) = \int_0^\infty \lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}-\beta\lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\phi\mathrm{e}^{-\phi\lambda}\mathrm{d}\lambda$$

$$
\begin{aligned}
&= \int_0^\infty \phi \mathrm{e}^{-\frac{y-\alpha}{\beta}} \lambda \mathrm{e}^{-\lambda\left(\phi+\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\phi \mathrm{e}^{-\frac{y-\alpha}{\beta}}}{\left(\phi+\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^2} \left(\phi+\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^2 \lambda \mathrm{e}^{-\lambda\left(\phi+\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)} \mathrm{d}\lambda \\
&= \frac{\phi \mathrm{e}^{-\frac{y-\alpha}{\beta}}}{\left(\phi+\beta \mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^2}
\end{aligned}
$$

When $\phi = \beta$, the resulting mixture has the form of a two parameter logistic density, Equation (B.26).
(Johnson and Kotz, 1970b, p. 3)

3.3.3 GAMMA

The following mixtures are all based on two parameter gamma densities, Equation (B.8), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized two parameter gamma distribution (closed under sampling).

Gaussian
The conditional density is a two parameter Gaussian, Equation (B.41).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\sqrt{\lambda}\mathrm{e}^{-\frac{\lambda(y-\mu)^2}{2}}}{\sqrt{2\pi}} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\beta\lambda}}{\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-\frac{1}{2}} \mathrm{e}^{-\lambda\left[\beta+\frac{(y-\mu)^2}{2}\right]}}{\sqrt{2\pi}\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\Gamma\left(\alpha+\frac{1}{2}\right)\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}\left[\beta+\frac{(y-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}}} \\
&\qquad\qquad \frac{\left[\beta+\frac{(y-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}} \lambda^{\alpha-\frac{1}{2}} \mathrm{e}^{-\lambda\left[\beta+\frac{(y-\mu)^2}{2}\right]}}{\Gamma\left(\alpha+\frac{1}{2}\right)} \mathrm{d}\lambda \\
&= \frac{\Gamma\left(\alpha+\frac{1}{2}\right)\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}\left[\beta+\frac{(y-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}}} \\
&= \frac{(2\beta)^\alpha}{B\left(\alpha,\frac{1}{2}\right)\left[2\beta+(y-\mu)^2\right]^{\alpha+\frac{1}{2}}}
\end{aligned}
$$

(Johnson and Kotz, 1970a, p. 88)

Poisson
The conditional density is a Poisson, Equation (B.45).

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^y \mathrm{e}^{-\lambda}}{y!} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha+y-1} \mathrm{e}^{-\lambda(\beta+1)}}{\Gamma(\alpha)\Gamma(y+1)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\Gamma(\alpha+y)\beta^\alpha}{\Gamma(\alpha)\Gamma(y)y(\beta+1)^{\alpha+y}} \frac{(\beta+1)^{\alpha+y}\lambda^{\alpha+y-1}\mathrm{e}^{-\lambda(\beta+1)}}{\Gamma(\alpha+y)} \mathrm{d}\lambda \\ &= \frac{\beta^\alpha}{B(\alpha,y)y(\beta+1)^{\alpha+y}} \end{aligned}$$

The resulting mixture has the form of a two parameter negative binomial distribution, Equation (B.48).
(Johnson and Kotz, 1969, pp. 124–125)

Gamma
The conditional density is a two parameter gamma, Equation (B.8).

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^\alpha y^{\alpha-1} \mathrm{e}^{-\lambda y}}{\Gamma(\alpha)} \frac{\phi^\psi \lambda^{\psi-1} \mathrm{e}^{-\lambda\phi}}{\Gamma(\psi)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{y^{\alpha-1}\phi^\psi \lambda^{\alpha+\psi-1} \mathrm{e}^{-\lambda(\phi+y)}}{\Gamma(\alpha)\Gamma(\psi)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\Gamma(\alpha+\psi)y^{\alpha-1}\phi^\psi}{\Gamma(\alpha)\Gamma(\psi)(\phi+y)^{\alpha+\psi}} \frac{(\phi+y)^{\alpha+\psi}\lambda^{\alpha+\psi-1}\mathrm{e}^{-\lambda(\phi+y)}}{\Gamma(\alpha+\psi)} \mathrm{d}\lambda \\ &= \frac{\phi^\psi y^{\alpha-1}}{B(\alpha,\psi)(\phi+y)^{\alpha+\psi}} \end{aligned}$$

The resulting mixture has the form of a two parameter beta distribution, Equation (B.34), where y is transformed by $\frac{\phi}{\phi+z}$ with Jacobian $\frac{\phi}{(\phi+z)^2}$.
(Johnson and Kotz, 1970a, p. 195)

Exponential
The conditional density is a one parameter exponential, Equation (B.5), resulting in the mixture density corresponding to a gamma Laplace transform.

$$\begin{aligned} f_{\mathrm{m}}(y) &= \int_0^\infty \lambda \mathrm{e}^{-\lambda y} \frac{\beta^\alpha \lambda^{\alpha-1} \mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\beta^\alpha \lambda^\alpha \mathrm{e}^{-\lambda[\beta+y]}}{\Gamma(\alpha)} \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\alpha\beta^\alpha}{[\beta+y]^{\alpha+1}} \frac{[\beta+y]^{\alpha+1}\lambda^\alpha \mathrm{e}^{-\lambda[\beta+y]}}{\Gamma(\alpha+1)} \mathrm{d}\lambda \end{aligned}$$

$$= \frac{\alpha\beta^\alpha}{[\beta+y]^{\alpha+1}}$$

The resulting mixture has the form of a two parameter Pareto density, Equation (B.23).
(Johnson and Kotz, 1970a, pp. 233–234; Cox and Oakes, 1984, pp. 19–20)

Generalization
The conditional density is chosen to have the form presented in Equation (3.2).

$$\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \kappa\lambda g(y)G(y)^{\kappa-1}\mathrm{e}^{-\lambda G(y)^\kappa}\frac{\beta^\alpha\lambda^{\alpha-1}\mathrm{e}^{-\lambda\beta}}{\Gamma(\alpha)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\kappa\beta^\alpha g(y)G(y)^{\kappa-1}\lambda^\alpha\mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}}{\Gamma(\alpha)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\alpha\kappa\beta^\alpha g(y)G(y)^{\kappa-1}}{[\beta+G(y)^\kappa]^{\alpha+1}}\frac{[\beta+G(y)^\kappa]^{\alpha+1}\lambda^\alpha\mathrm{e}^{-\lambda[\beta+G(y)^\kappa]}}{\Gamma(\alpha+1)}\mathrm{d}\lambda \\
&= \frac{\alpha\kappa\beta^\alpha g(y)G(y)^{\kappa-1}}{[\beta+G(y)^\kappa]^{\alpha+1}}
\end{aligned}$$

When the conditional density is chosen to be a two parameter Weibull, Equation (B.12), the resulting mixture has the form of a three parameter Burr density, Equation (B.35).
(Johnson and Kotz, 1970a, p. 266)

Gumbel
The conditional density is a two parameter Gumbel, Equation (B.15), where the location is re-parameterized as $\alpha+\beta\ln[\lambda\beta]$ and the scale is set equal to β.

$$\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}-\beta\lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\frac{\phi^\psi\lambda^{\psi-1}\mathrm{e}^{-\lambda\phi}}{\Gamma(\psi)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\phi^\psi\mathrm{e}^{-\frac{y-\alpha}{\beta}}\lambda^\psi\mathrm{e}^{-\lambda\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}}{\Gamma(\psi)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\psi\phi^\psi\mathrm{e}^{-\frac{y-\alpha}{\beta}}}{\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^{\psi+1}}\frac{\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^{\psi+1}\lambda^\psi\mathrm{e}^{-\lambda\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}}{\Gamma(\psi+1)}\mathrm{d}\lambda \\
&= \frac{\psi\phi^\psi\mathrm{e}^{-\frac{y-\alpha}{\beta}}}{\left(\phi+\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^{\psi+1}}
\end{aligned}$$

The resulting mixture has the form of a four parameter generalized logistic density, Equation (B.30).
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

Laplace

The conditional density is a two parameter Laplace, Equation (B.39).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda e^{-\lambda|y-\mu|}}{2} \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} d\lambda \\
&= \int_0^\infty \frac{\beta^\alpha \lambda^\alpha e^{-\lambda[\beta+|y-\mu|]}}{2\Gamma(\alpha)} d\lambda \\
&= \int_0^\infty \frac{\alpha\beta^\alpha}{2[\beta+|y-\mu|]^{\alpha+1}} \frac{[\beta+|y-\mu|]^{\alpha+1}\lambda^\alpha e^{-\lambda[\beta+|y-\mu|]}}{\Gamma(\alpha+1)} d\lambda \\
&= \frac{\alpha\beta^\alpha}{2[\beta+|y-\mu|]^{\alpha+1}}
\end{aligned}
$$

(Johnson and Kotz, 1970b, p. 32)

Simplex

The conditional density is a two parameter simplex, Equation (B.40), where $\lambda = \frac{1}{\sigma^2}$.

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \sqrt{\frac{\lambda}{2\pi y^3(1-y)^3}} e^{-\lambda\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}} \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} d\lambda \\
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-\frac{1}{2}} e^{-\lambda\left[\beta+\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}\right]}}{\sqrt{2\pi y^3(1-y)^3}\,\Gamma(\alpha)} d\lambda \\
&= \int_0^\infty \frac{\Gamma(\alpha+\frac{1}{2})\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi y^3(1-y)^3}\left[\beta+\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}\right]^{\alpha+\frac{1}{2}}} \\
&\qquad \frac{\left[\beta+\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}\right]^{\alpha+\frac{1}{2}} \lambda^{\alpha-\frac{1}{2}} e^{-\lambda\left[\beta+\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}\right]}}{\Gamma(\alpha+\frac{1}{2})} d\lambda \\
&= \frac{\Gamma(\alpha+\frac{1}{2})\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi y^3(1-y)^3}\left[\beta+\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}\right]^{\alpha+\frac{1}{2}}} \\
&= \frac{(2\beta)^\alpha}{B\left(\alpha,\frac{1}{2}\right)\sqrt{y^3(1-y)^3}\left[2\beta+\frac{(y-\mu)^2}{y(1-y)\mu^2(1-\mu)^2}\right]^{\alpha+\frac{1}{2}}}
\end{aligned}
$$

Lognormal

The conditional density is a three parameter lognormal, Equation (B.44), where $\lambda = \frac{1}{\sigma^2}$.

$$
f_{\mathrm{m}}(y) = \int_0^\infty \frac{\sqrt{\lambda}e^{-\frac{\lambda}{2}(\ln[y-\theta]-\mu)^2}}{(y-\theta)\sqrt{2\pi}} \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} d\lambda
$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-\frac{1}{2}} \mathrm{e}^{-\lambda\left[\beta+\frac{(\ln[y-\theta]-\mu)^2}{2}\right]}}{(y-\theta)\sqrt{2\pi}\Gamma(\alpha)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\Gamma(\alpha+\frac{1}{2})\beta^\alpha}{\Gamma(\alpha)(y-\theta)\sqrt{2\pi}\left[\beta+\frac{(\ln[y-\theta]-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}}} \\
&\qquad \frac{\left[\beta+\frac{(\ln[y-\theta]-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}} \lambda^{\alpha-\frac{1}{2}} \mathrm{e}^{-\lambda\left[\beta+\frac{(\ln[y-\theta]-\mu)^2}{2}\right]}}{\Gamma(\alpha+\frac{1}{2})} \mathrm{d}\lambda \\
&= \frac{\Gamma(\alpha+\frac{1}{2})\beta^\alpha}{\Gamma(\alpha)(y-\theta)\sqrt{2\pi}\left[\beta+\frac{(\ln[y-\theta]-\mu)^2}{2}\right]^{\alpha+\frac{1}{2}}} \\
&= \frac{(2\beta)^\alpha}{B\left(\alpha,\frac{1}{2}\right)(y-\theta)\left[2\beta+(\ln[y-\theta]-\mu)^2\right]^{\alpha+\frac{1}{2}}}
\end{aligned}
$$

3.3.4 INVERSE GAUSSIAN

The following mixtures are all based on two parameter inverse Gaussian densities, Equation (B.9), as mixing distribution. When combined with a gamma, the closed form mixture is obtained by integrating out a re-parameterized three parameter generalized inverse Gaussian distribution. On the other hand when combined with an exponential density, a density of the form presented in Equation (3.2), a Gumbel, or Laplace density, the closed form mixture is obtained by integrating out a re-parameterized one parameter random walk distribution, Equation (B.11).

Gamma

The conditional density is a two parameter gamma, Equation (B.8).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^\alpha y^{\alpha-1}\mathrm{e}^{-\lambda y}}{\Gamma(\alpha)} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi}} \frac{y^{\alpha-1}\mathrm{e}^{\sqrt{\psi\phi}}}{\Gamma(\alpha)} \lambda^{\alpha-\frac{3}{2}} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi}} \frac{2y^{\alpha-1}\mathrm{e}^{\sqrt{\psi\phi}}\left(\frac{\psi}{\phi+2y}\right)^{\frac{2\alpha-1}{4}} K\left(\sqrt{\psi(\phi+2y)}, \left|\alpha-\frac{1}{2}\right|\right)}{\Gamma(\alpha)} \\
&\qquad \frac{\lambda^{\alpha-\frac{3}{2}} \mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2y}\right)^{\frac{2\alpha-1}{4}} K\left(\sqrt{\psi(\phi+2y)}, \left|\alpha-\frac{1}{2}\right|\right)} \mathrm{d}\lambda \\
&= \sqrt{\frac{\psi}{2\pi}} \frac{2y^{\alpha-1}\mathrm{e}^{\sqrt{\psi\phi}}\left(\frac{\psi}{\phi+2y}\right)^{\frac{2\alpha-1}{4}} K\left(\sqrt{\psi(\phi+2y)}, \left|\alpha-\frac{1}{2}\right|\right)}{\Gamma(\alpha)}
\end{aligned}
$$

$$= \frac{K\left(\sqrt{\psi(\phi+2y)}, \left|\alpha-\frac{1}{2}\right|\right)}{\Gamma(\alpha)} \sqrt{\frac{2}{\pi}\left(\frac{\psi}{\phi+2y}\right)^{\alpha}\sqrt{\psi(\phi+2y)}}\, y^{\alpha-1} \mathrm{e}^{\sqrt{\psi\phi}}$$

Exponential

The conditional density is a one parameter exponential, Equation (B.5), resulting in the mixture density corresponding to an inverse Gaussian Laplace transform.

$$\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda \mathrm{e}^{-\lambda y} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma \\
&= \int_0^\infty \frac{\psi}{a} \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{a}{2\pi\gamma}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \mathrm{d}\gamma \\
&= \frac{\psi}{a} \mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \sqrt{\frac{\psi}{\phi+2y}} \mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi(\phi+2y)}}
\end{aligned}$$

Where $a = \sqrt{\psi(\phi+2y)}$ and $\gamma = \frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda = \frac{\psi}{a}\mathrm{d}\gamma$.

Generalization

The conditional density is chosen to have the form presented in Equation (3.2).

$$\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \kappa\lambda g(y) G(y)^{\kappa-1} \mathrm{e}^{-\lambda G(y)^\kappa} \sqrt{\frac{\psi}{2\pi\lambda^3}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \kappa g(y) G(y)^{\kappa-1} \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}(\phi+2G(y)^\kappa)-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \kappa g(y) G(y)^{\kappa-1} \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}} \mathrm{d}\lambda \\
&= \int_0^\infty \kappa g(y) G(y)^{\kappa-1} \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{\sqrt{\psi\phi}-a+a-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \kappa g(y) G(y)^{\kappa-1} \mathrm{e}^{\sqrt{\psi\phi}-a} \sqrt{\frac{\psi}{2\pi\lambda}} \mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)} \frac{\psi}{a} \mathrm{d}\gamma
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\psi}{a}\kappa g(y)G(y)^{\kappa-1}\mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{a}{2\pi\gamma}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\mathrm{d}\gamma \\
&= \kappa g(y)G(y)^{\kappa-1}\frac{\psi}{a}\mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \kappa g(y)G(y)^{\kappa-1}\sqrt{\frac{\psi}{\phi+2G(y)^\kappa}}\mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi[\phi+2G(y)^\kappa]}}
\end{aligned}
$$

Where $a=\sqrt{\psi\left[\phi+2G(y)^\kappa\right]}$ and $\gamma=\frac{a\lambda}{\psi}$ with Jacobian $\mathrm{d}\lambda=\frac{\psi}{a}\mathrm{d}\gamma$.

Gumbel
The conditional density is a two parameter Gumbel, Equation (B.15), where the location is re-parameterized as $\alpha+\beta\ln[\lambda\beta]$ and the scale is set equal to β.

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}-\beta\lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{-\frac{y-\alpha}{\beta}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{-\frac{y-\alpha}{\beta}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{-\frac{y-\alpha}{\beta}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{-\frac{y-\alpha}{\beta}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\frac{\psi}{a}\mathrm{d}\gamma \\
&= \int_0^\infty \frac{\psi}{a}\mathrm{e}^{\sqrt{\psi\phi}-a-\frac{y-\alpha}{\beta}}\sqrt{\frac{a}{2\pi\gamma}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\mathrm{d}\gamma \\
&= \frac{\psi}{a}\mathrm{e}^{\sqrt{\psi\phi}-a-\frac{y-\alpha}{\beta}} \\
&= \sqrt{\frac{\psi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{y-\alpha}{\beta}-\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)}}
\end{aligned}
$$

Laplace
The conditional density is a two parameter Laplace, Equation (B.39).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda\mathrm{e}^{-\lambda|y-\mu|}}{2}\sqrt{\frac{\psi}{2\pi\lambda^3}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi\lambda+\frac{\psi}{\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{1}{2}\sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{\lambda}{2}[\phi+2|y-\mu|]-\frac{\psi}{2\lambda}}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{1}{2}\sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a^2\lambda}{2\psi}-\frac{\psi}{2\lambda}}\mathrm{d}\lambda
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{2}\sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{a}{2}\left(\frac{a\lambda}{\psi}+\frac{\psi}{a\lambda}\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{1}{2}\mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{\psi}{2\pi\lambda}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\frac{\psi}{a}\mathrm{d}\gamma \\
&= \int_0^\infty \frac{\psi}{2a}\mathrm{e}^{\sqrt{\psi\phi}-a}\sqrt{\frac{a}{2\pi\gamma}}\mathrm{e}^{a-\frac{a}{2}\left(\gamma+\frac{1}{\gamma}\right)}\mathrm{d}\gamma \\
&= \frac{\psi}{2a}\mathrm{e}^{\sqrt{\psi\phi}-a} \\
&= \frac{\sqrt{\psi}\mathrm{e}^{\sqrt{\psi\phi}-\sqrt{\psi[\phi+2|y-\mu|]}}}{2\sqrt{\phi+2|y-\mu|}}
\end{aligned}
$$

3.3.5 GENERALIZED INVERSE GAUSSIAN

The following mixtures are all based on three parameter generalized inverse Gaussian densities, Equation (B.10), as mixing distribution. Each closed form mixture is obtained by integrating out a re-parameterized three parameter generalized inverse Gaussian distribution (closed under sampling).

Poisson
The conditional density is a Poisson, Equation (B.45).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^y \mathrm{e}^{-\lambda}}{y!}\frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^{\epsilon+y-1}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2)-\frac{\psi}{2\lambda}}}{2\Gamma(y+1)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\left(\frac{\psi}{\phi+2}\right)^{\frac{\epsilon+y}{2}}K\left(\sqrt{\psi(\phi+2)},|\epsilon+y|\right)}{\Gamma(y+1)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)} \\
&\qquad \frac{\lambda^{\epsilon+y-1}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2}\right)^{\frac{\epsilon+y}{2}}K\left(\sqrt{\psi(\phi+2)},|\epsilon+y|\right)}\mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi(\phi+2)},|\epsilon+y|\right)}{K\left(\sqrt{\psi\phi},|\epsilon|\right)}\frac{\psi^{\frac{y}{2}}\phi^{\frac{\epsilon}{2}}}{\Gamma(y+1)(\phi+2)^{\frac{\epsilon+y}{2}}}
\end{aligned}
$$

Gamma
The conditional density is a two parameter gamma, Equation (B.8).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda^\alpha y^{\alpha-1} \mathrm{e}^{-\lambda y}}{\Gamma(\alpha)} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{y^{\alpha-1}\lambda^{\epsilon+\alpha-1}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{\Gamma(\alpha) 2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{y^{\alpha-1}\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon+\alpha}{2}} K\left(\sqrt{\psi(\phi+2y)},|\epsilon+\alpha|\right)}{\Gamma(\alpha)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \\
&\qquad\qquad \frac{\lambda^{\epsilon+\alpha-1}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon+\alpha}{2}} K\left(\sqrt{\psi(\phi+2y)},|\epsilon+\alpha|\right)} \mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi(\phi+2y)},|\epsilon+\alpha|\right)}{K\left(\sqrt{\psi\phi},|\epsilon|\right)\Gamma(\alpha)} \sqrt{\frac{\psi^\alpha\phi^\epsilon}{(\phi+2y)^{\epsilon+\alpha}} y^{\alpha-1}}
\end{aligned}
$$

Exponential

The conditional density is a one parameter exponential, Equation (B.5), resulting in the mixture density corresponding to an generalized inverse Gaussian Laplace transform.

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda \mathrm{e}^{-\lambda y} \frac{\lambda^{\epsilon-1} \mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi},|\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{K\left(\sqrt{\psi(\phi+2y)},|\epsilon+1|\right)\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon+1}{2}}}{K\left(\sqrt{\psi\phi},|\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \\
&\qquad\qquad \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2y)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2y}\right)^{\frac{\epsilon+1}{2}} K\left(\sqrt{\psi(\phi+2y)},|\epsilon+1|\right)} \mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi(\phi+2y)},|\epsilon+1|\right)}{K\left(\sqrt{\psi\phi},|\epsilon|\right)} \frac{\sqrt{\psi}\phi^{\frac{\epsilon}{2}}}{(\phi+2y)^{\frac{\epsilon+1}{2}}}
\end{aligned}
$$

Generalization

The conditional density is chosen to have the form presented in Equation (3.2).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda\kappa g(y)G(y)^{\kappa-1}\mathrm{e}^{-\lambda G(y)^\kappa}\frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\kappa g(y)G(y)^{\kappa-1}\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}[\phi+2G(y)^\kappa]-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{K\left(\sqrt{\psi[\phi+2G(y)^\kappa]},|\epsilon+1|\right)\left(\frac{\psi}{\phi+2G(y)^\kappa}\right)^{\frac{\epsilon+1}{2}}\kappa g(y)G(y)^{\kappa-1}}{K\left(\sqrt{\psi\phi},|\epsilon|\right)\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}} \\
&\qquad \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}(\phi+2G(y)^\kappa)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2G(y)^\kappa}\right)^{\frac{\epsilon+1}{2}}K\left(\sqrt{\psi[\phi+2G(y)^\kappa]},|\epsilon+1|\right)}\mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi[\phi+2G(y)^\kappa]},|\epsilon+1|\right)}{K\left(\sqrt{\psi\phi},|\epsilon|\right)}\frac{\sqrt{\psi}\phi^{\frac{\epsilon}{2}}}{[\phi+2G(y)^\kappa]^{\frac{\epsilon+1}{2}}}\kappa g(y)G(y)^{\kappa-1}
\end{aligned}
$$

Gumbel

The conditional density is a two parameter Gumbel, Equation (B.15), where the location is re-parameterized as $\alpha+\beta\ln[\lambda\beta]$ and the scale is set equal to β.

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}-\beta\lambda\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\mathrm{e}^{-\frac{y-\alpha}{\beta}}\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)}\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\mathrm{e}^{-\frac{y-\alpha}{\beta}}\left(\frac{\psi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\right)^{\frac{\epsilon+1}{2}}K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)},|\epsilon+1|\right)}{\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}}K\left(\sqrt{\psi\phi},|\epsilon|\right)} \\
&\qquad \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}}\right)^{\frac{\epsilon+1}{2}}K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)},|\epsilon+1|\right)}\mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)},|\epsilon+1|\right)}{K\left(\sqrt{\psi\phi},|\epsilon|\right)}\frac{\sqrt{\psi}\phi^{\frac{\epsilon}{2}}\mathrm{e}^{-\frac{y-\alpha}{\beta}}}{\left(\phi+2\beta\mathrm{e}^{-\frac{y-\alpha}{\beta}}\right)^{\frac{\epsilon+1}{2}}}
\end{aligned}
$$

Laplace
The conditional density is a two parameter Laplace, Equation (B.39).

$$
\begin{aligned}
f_{\mathrm{m}}(y) &= \int_0^\infty \frac{\lambda \mathrm{e}^{-\lambda|y-\mu|}}{2} \frac{\lambda^{\epsilon-1}\mathrm{e}^{-\frac{\lambda\phi}{2}-\frac{\psi}{2\lambda}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}[\phi+2|y-\mu|]-\frac{\psi}{2\lambda}}}{4\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \mathrm{d}\lambda \\
&= \int_0^\infty \frac{\left(\frac{\psi}{\phi+2|y-\mu|}\right)^{\frac{\epsilon+1}{2}} K\left(\sqrt{\psi[\phi+2|y-\mu|]}, |\epsilon+1|\right)}{2\left(\frac{\psi}{\phi}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \\
&\qquad\qquad \frac{\lambda^{\epsilon}\mathrm{e}^{-\frac{\lambda}{2}[\phi+2|y-\mu|]-\frac{\psi}{2\lambda}}}{\left(\frac{\psi}{\phi+2|y-\mu|}\right)^{\frac{\epsilon+1}{2}} K\left(\sqrt{\psi[\phi+2|y-\mu|]}, |\epsilon+1|\right)} \mathrm{d}\lambda \\
&= \frac{K\left(\sqrt{\psi[\phi+2|y-\mu|]}, |\epsilon+1|\right)}{K\left(\sqrt{\psi\phi}, |\epsilon|\right)} \frac{\sqrt{\psi}\phi^{\frac{\epsilon}{2}}}{2[\phi+2|y-\mu|]^{\frac{\epsilon+1}{2}}}
\end{aligned}
$$

3.4 Further reading

4
Multiple response & cluster models

In this chapter, the multivariate models introduced are based on the Archimedean copulas and multivariate Laplace transforms described in Subsection 2.2.2. Section 4.1 describes the method used to take the appropriate differences of a multivariate cdf in order to obtain the multivariate probability mass function.

As a first approach, the method to obtain the copula corresponding to known multivariate distributions is discussed in Section 4.2. Unfortunately, this approach is limited as only few multivariate distributions are available and certain of these are based on the same multivariate dependence structure (copula).

From Equations (2.21) and (2.24), we see that these Archimedean copulas are permutation symmetric in their margins. This corresponds to a frailty model in a survival context. Although exchangeability can be very useful, for example in the case of clustered data, methods of combining Archimedean (survival) copulas together will be discussed in Section 4.3,in order to obtain more specific types of dependencies. For instance, this will allow introduction of features such as asymmetries among margins or higher concordance among certain margins than others which can be useful to model multiple response data.

At the end of Subsection 4.3.5, these techniques are adapted to induce different types of structured dependencies required in the case of longitudinal data. But this field will only be briefly discussed because Chapter 5, dedicated to longitudinal data models, will cover more extensively these types of responses.

4.1 Probabilities in m dimensions

Because multivariate cdfs are not linear in their parameters, numerical optimization will be used to obtain (simultaneously) maximum likelihood estimates of the dependence parameters as well as the (nonlinear) regression parameters in each margins. Obviously for this, the multivariate likelihood must be constructed. This will be illustrated for the special case where the multivariate cdf is a copula.

Recall from Theorem 2.14 and Corollary 2.15 that (survival) copulas are just dependency structures linking the margins together. Hence, a specific set of univariate cdfs must be chosen for the margins in order to define a multivariate cdf. Because the responses under study are nominal or ordinal, the margins are chosen as univariate cdfs corresponding to one of the parametrizations mentioned in Subsections 1.3.1 to 1.3.6. But for ordinal data, the joint probabilities are needed for

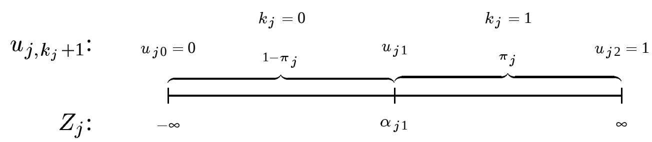

Fig. 4.1. The cumulative probabilities (u_{j,k_j}) and probabilities (π_{k_j}) for binary margin j corresponding to an underlying continuous variable Z_j with cut point α_{j1}.

the multivariate likelihood, not the joint cumulative probabilities. This requires the multivariate probability mass function which can be obtained by taking the appropriate differences of the multivariate cdf.

For simplicity, a bivariate copula with binary margins will first be considered. The cumulative probabilities (u_{jk_j}) and probabilities (π_{k_j}) are shown in Figure 4.1 for one of the two margins. Note that u_{j,k_j+1} is the cumulative probability of margin j ($j = 1,2$) observed in category k_j ($k_j = 0,1$), so that u_{j0} is equal to zero whereas u_{j2} is equal to one.

In some cases, it is reasonable to assume that the nominal or ordinal responses observed are the result of the discretization of a continuous multivariate family $F_{\mathrm{M}} \in F\,(F_1,\ldots,F_m)$. In other words, the responses are following an underlying continuous variable or latent variable $\mathbf{Z} \sim F_{\mathrm{M}}$. Sometimes, it is also reasonable further to assume that each of the univariate margins is as well following a latent variable $Z_j \sim F_j$. This latent variable Z_j will then have cut points α_{j,k_j+1} for $k_j = 0,\ldots,(\mathtt{nc}-2)$ where `nc` denotes the number of observable categories.

The bivariate cumulative probabilities are shown in Figure 4.2a. Note that there are four joint cumulative probabilities because each margin has two possible outcomes. These are $C_{\mathrm{B}}(u_{1k_1},u_{2k_2})$, $C_{\mathrm{B}}(u_{1k_1},1) = u_{1k_1}$, $C_{\mathrm{B}}(1,u_{2k_2}) = u_{2k_2}$, and $C_{\mathrm{B}}(1,1) = 1$. Recall from Definition 2.8 that $C_{\mathrm{B}}(u_{1k_1},0) = C_{\mathrm{B}}(0,u_{2k_2}) = 0$.

Now, the bivariate probabilities can be obtained by taking the appropriate differences. It follows from Figure 4.2b that the four joint probabilities are

$$
\begin{aligned}
\Pr(Y_1 = 0, Y_2 = 0) &= C_{\mathrm{B}}(u_{1k_1},u_{2k_2}) - C_{\mathrm{B}}(u_{1k_1},0) - C_{\mathrm{B}}(0,u_{2k_2}) + C_{\mathrm{B}}(0,0)\\
&= \pi_{00}\\
\Pr(Y_1 = 0, Y_2 = 1) &= C_{\mathrm{B}}(u_{1k_1},1) - C_{\mathrm{B}}(0,1) - C_{\mathrm{B}}(u_{1k_1},u_{2k_2}) + C_{\mathrm{B}}(0,u_{2k_2})\\
&= u_{1k_1} - \pi_{00}\\
&= \pi_{01}\\
\Pr(Y_1 = 1, Y_2 = 0) &= C_{\mathrm{B}}(1,u_{2k_2}) - C_{\mathrm{B}}(1,0) - C_{\mathrm{B}}(u_{1k_1},u_{2k_2}) + C_{\mathrm{B}}(u_{1k_1},0)\\
&= u_{2k_2} - \pi_{00}\\
&= \pi_{10}\\
\Pr(Y_1 = 1, Y_2 = 1) &= C_{\mathrm{B}}(1,1) - C_{\mathrm{B}}(1,u_{2k_2}) - C_{\mathrm{B}}(u_{1k_1},1) + C_{\mathrm{B}}(u_{1k_1},u_{2k_2})\\
&= 1 - u_{2k_2} - u_{1k_1} + \pi_{00}\\
&= \pi_{11}
\end{aligned}
$$

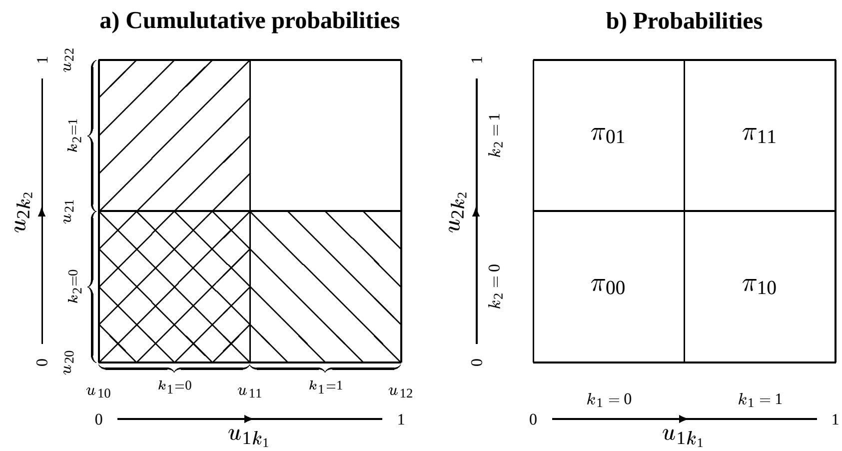

Fig. 4.2. The joint cumulative probabilities (a) and joint probabilities (b) of a bivariate copula with binary margins.

Extension to more than two categories is straightforward because the appropriate differences always have the same form in the bivariate case.

$$\begin{aligned}\Pr(Y_1 = k_1, Y_2 = k_2) = C_{\mathrm{B}}(u_{1,k_1+1}, u_{2,k_2+1}) - C_{\mathrm{B}}(u_{1,k_1+1}, u_{2,k_2}) \\ - C_{\mathrm{B}}(u_{1,k_1}, u_{2,k_2+1}) + C_{\mathrm{B}}(u_{1,k_1}, u_{2,k_2})\end{aligned}$$

The bivariate probability mass function for `nc` categories can be written as

$$\Pr(Y_1 = k_1, Y_2 = k_2) = \sum_{s_1=k_1}^{k_1+1} \sum_{s_2=k_2}^{k_2+1} (-1)^{2-\Sigma_{r=1}^{2}(k_r - s_r)} C_{\mathrm{B}}(u_{1,s_1}, u_{2,s_2})$$

where $k_j = 0, \ldots, (\mathtt{nc} - 1)$ for $j = 1, 2$.

Finally, the multivariate probability mass function is

$$\Pr(Y_1 = k_1, \ldots, Y_m = k_m) = \sum_{s_1=k_1}^{k_1+1} \cdots \sum_{s_m=k_m}^{k_m+1} (-1)^{m-\Sigma_{r=1}^{m}(k_r - s_r)} C_{\mathrm{M}}(u_{1,s_1}, \ldots, u_{m,s_m})$$

where the number of differences has been adapted accordingly to the dimension of the copula.

As mentioned earlier, this same procedure works for any multivariate cdf, not just copulas. In the case of multivariate survival functions, the technique must slightly be adapted. Indeed, differences must in such a case be taken between a given category and a combination of higher ones (no longer a combination of lower categories).

4.2 Remapping margins of multivariate distributions

The first approach to obtain flexible families of multivariate distributions is to extend already existing multivariate distributions. A method found in the literature consists in remapping the univariate marginal distributions of a known multivariate distribution.

This can be achieved using Corollary 2.15.

$$\begin{aligned} \widehat{F}_{\mathrm{M}}(\mathbf{y}) &= F_{\mathrm{M}}\left(F^{-1}\left(\widetilde{F}_1(y_1)\right),\ldots,F^{-1}\left(\widetilde{F}_m(y_m)\right)\right) \\ &= C_{\mathrm{M}}\left(\widetilde{F}_1(y_1),\ldots,\widetilde{F}_m(y_m)\right) \end{aligned} \tag{4.1}$$

A new multivariate distribution $\widehat{F}_{\mathrm{M}}$ is obtained from a known multivariate distribution F_{M} by inserting remapped margins of the form

$$F^{-1}\left(\widetilde{F}(y)\right)$$

where F^{-1} is the quantile function corresponding to the univariate case of F_{M} and $\widetilde{F}$ is an arbitrarily chosen cdf different from F.

By this technique, new multivariate distributions are generated through the intermediate step of copulas. This implies that the marginal univariate distributions F chosen defines the resulting multivariate distribution. Hence, many general families of multivariate distributions can be produced where the family is determined by the copula or the known multivariate distribution with uniformly $[0,1]$ standardized margins.

In situations where it is reasonable to assume that a nominal or ordinal response is following an underlying continuous variable (see Section 4.1 and Subsection 5.3.1), the multivariate density c_{M} corresponding to a (survival) copula C_{M} will be of interest.

Note that the assumption of an underlying continuous variable will generally not be made unless clearly stated. Therefore, the probabilities will usually be obtained from the probability mass function which takes the appropriate differences of the (survival) copula as is done in Section 4.1. But in the particular case where this assumption is made, the multivariate density — obtained by taking derivatives of the (survival) copula with respect to each of the univariate margins — is used.

From Equation (4.1), the multivariate density $\widehat{f}_{\mathrm{M}}$ constructed by remapping the margins of a multivariate distribution has as general form

$$\widehat{f}_{\mathrm{M}}(\mathbf{y}) = c_{\mathrm{M}}\left(\widetilde{F}_1(y_1),\ldots,\widetilde{F}_m(y_m)\right) f_1(y_1)\ldots f_m(y_m)$$

where c_{M} must be multiplied by the appropriate Jacobians corresponding to the transformations applied to its respective margins (Joe, 1997, p. 299; Song, 2000). As the transformation applied to margin i of c_{M} is a univariate cdf F_i, the Jacobian is the corresponding univariate density f_i. Note that care must be taken if starting out from a multivariate density because the joint function corresponding to the copula must also include the Jacobians of the quantile transformations.

A first example would be to consider the Gaussian copula obtained from the multivariate Gaussian distribution, Equation (B.58), and univariate Gaussian quantile functions, the inverse of the standardized Gaussian function given in equation (B.4). The one-parameter family of copulas obtained in the bivariate case is discussed by Joe (1997, pp. 140–141). The general multivariate case is described by Song (2000) where the full covariance matrix of dependence parameters is available. This paper mainly discusses the case where members of the exponential dispersion family are inserted into the Gaussian copula generating multivariate dispersion distributions.

Multivariate elliptically-contoured distributions could also be used to generate families of copulas. The two most well known are the multivariate power exponential and the multivariate Student t distributions respectively given in Equations (B.56) and (B.57).

The multivariate power exponential distribution was obtained by Gómez *et al.* (1998) as a generalization of the univariate power exponential distribution. It includes the multivariate Gaussian, the multivariate Laplace, and the multivariate uniform distributions as special cases. A full covariance matrix of dependence parameters is available although care must be taken if independence is wanted among certain observations (see Lindsey, 1999, pp. 54–55 for further details).

The multivariate Student t distribution includes the multivariate Cauchy and multivariate Gaussian distributions as special cases. Hence, a full covariance matrix of dependence parameters is also available.

Other multivariate distributions which will be of interest are the multivariate Burr, multivariate logistic, multivariate generalized logistic, multivariate Pareto, and multivariate exponential distributions respectively given in Equations (B.49), (B.50), (B.52), (B.54), and (B.55). Note that these multivariate distribution have only one dependence parameter but that the multivariate cdf or survival function has a closed form.

Finally, multivariate distributions obtained dynamically can also be used in this context. Indeed, Jones (1993) shows how the multivariate Gaussian distribution, Equation (B.58), can be built through dynamically updated equations. Thus, the multivariate dynamic distributions provided in Chapter 5 will also be suitable for margin remapping.

4.3 Archimedean copulas

From the theory presented in Subsection 2.2.2, many bivariate and multivariate Archimedean (survival) copulas can easily be obtained from strictly decreasing and convex functions on $[0,1]$. But a parameter measuring the strength of dependence among the margins is still missing.

Two straightforward methods exist for adding a dependence parameter to these generators creating a large class of one-parameter families of Archimedean (survival) copulas.

THEOREM 4.1 *[Nelsen, 1999, p. 114 (proof)]*

Let $\psi(t)$ be in Ω the set of continuous strictly decreasing convex functions ψ from $[0,1]$ to $[0,\infty)$ with $\psi(1)=0$. And let α and β be positive real numbers. Then let

$$\psi_{\alpha,1}(t)=\psi(t^{\alpha}) \quad \text{and} \quad \psi_{1,\beta}(t)=[\psi(t)]^{\beta} \tag{4.2}$$

(1) If $\beta \geq 1$ then $\psi_{1,\beta}(t)$ is an element of Ω.
(2) If α is in $(0,1]$ then $\psi_{\alpha,1}(t)$ is an element of Ω.
(3) If $\psi(t)$ is twice differentiable and $t\psi'(t)$ is nondecreasing on $(0,1)$ then $\psi_{\alpha,1}(t)$ is an element of Ω for all $\alpha > 0$.

The one-parameter Archimedean (survival) copulas obtained in this way will respectively be referred to as alpha-generated and beta-generated families. □

Appendix C.1 contains a list of one-parameter families of Archimedean (survival) copulas. A total of 22 families are presented in their bivariate and multivariate form, along with their respective generator(s) or Laplace transform(s), parameter range, boundary and independence cases, and plots of a bivariate dependency for certain choices of margins.

An interesting property can be seen for alpha-generated and beta-generated families of Archimedean (survival) copulas.

THEOREM 4.2 *[Nelsen, 1999, p. 115 (proof)]*

Let $\psi(t)$ be in Ω and let $\psi_{\alpha,1}(t)$ and $\psi_{1,\beta}(t)$ be defined as in Theorem 4.1. Further assume that $\psi_{\alpha,1}(t)$ and $\psi_{1,\beta}(t)$ respectively generate copulas $C_{\alpha,1}(\mathbf{u})$ and $C_{1,\beta}(\mathbf{u})$ where $\beta \geq 1$ and α is an element of a subset of $(0,\infty)$ which includes $(0,1]$.

(1) If $\psi(t)$ is continuously differentiable and $\psi'(1) \neq 0$ then

$$C_{0,1}(\mathbf{u}) = \lim_{\alpha\to 0^+} C_{\alpha,1}(\mathbf{u}) = \prod_{j=1}^{m} u_j$$

(2)

$$C_{1\infty}(\mathbf{u}) = \lim_{\beta\to\infty} C_{1,\beta}(\mathbf{u}) = F_U$$

Hence, all the alpha-generated families include independence as a limiting case, while all the beta-generated families include the Fréchet upper bound as a limiting case. This can indeed be seen for the relevant Archimedean (survival) copulas presented in Appendix C.1. □

In the bivariate case, the parameter range can sometimes be extended to include a negative dependence. Such an extension is known for two families of bivariate Archimedean (survival) copulas of Appendix C.1 and can be found in

Appendix C.4 with plots of a negative bivariate dependency for certain choices of margins.

Hence, this section will now concentrate on extensions introducing additional parameters to Archimedean (survival) copulas and specific parametrizations of these to model certain types of dependent observations.

4.3.1 ADDITIONAL PROPERTIES

The following series of definitions are generally used to obtain specific properties or characteristics of (bivariate) Archimedean (survival) copulas. Certain of the properties deduced will be of particular interest in the following subsections in order to obtain proper higher order dependency models.

As a starting point, we shall look at three special cases of Archimedean (survival) copulas. All three, respectively the Fréchet lower bound (F_L), independence, and the Fréchet upper bound (F_U), are functions obtained when the appropriate one-sided limit of the dependence parameter approaches an end point of its interval (see Theorem 2.16 for the Fréchet bounds).

(1) THEOREM 4.3 *[Joe, 1997, pp. 61–62 (proof); Nelsen, 1999, p. 42 (proof from Sklar)]*

A necessary and sufficient condition for the F_L to be a multivariate cdf (or survival function) is that either

- $\sum_j u_j \leq 1$ whenever $0 < u_j < 1$ or
- $\sum_j u_j \geq m-1$ whenever $0 < u_j < 1$,

for $j = 1, \ldots, m$. Hence depending on $\mathbf{u}$, there exists an Archimedean (survival) copula $C_{\mathrm{M}}(\mathbf{u})$ equal to the F_L, but in general the F_L is never a copula for $m > 2$. □

(2) DEFINITION 4.4 *[Nelsen, 1999, pp. 9, 42, and 112]*

An important Archimedean (survival) copula occurs at independence and will be called the product copula P. □

(3) DEFINITION 4.5 *[Joe, 1997, p. 58 (proof); Nelsen, 1999, pp. 9, 42, and 112]*

The F_U is a multivariate cdf or survival function and therefore a (survival) copula but not Archimedean. □

Figure 4.3 shows the contour plots of the bivariate cumulative probabilities corresponding to the Fréchet lower bound, independence, and the Fréchet upper bound.

Now, methods to obtain the boundaries and independence are required. The following two techniques can generally be used, although sometimes l'Hôpital's rule (Salas and Hille, 1990, pp. 591 and 597) might also be required.

(1) THEOREM 4.6 *[Genest and MacKay, 1986a; Nelsen, 1999, p. 112 (proof)]*

Let $C_{\mathrm{M}}(\mathbf{u};\theta)$ for $\theta \in \Theta$ be a family of Archimedean (survival) copulas with differentiable generators $\psi_\theta(s)$ in Ω. Then the limit of $C_{\mathrm{M}}(\mathbf{u};\theta)$ is equal to the F_L or independence if and only if there exists a function ψ in Ω such that for all s_1, s_2 in $(0,1)$,

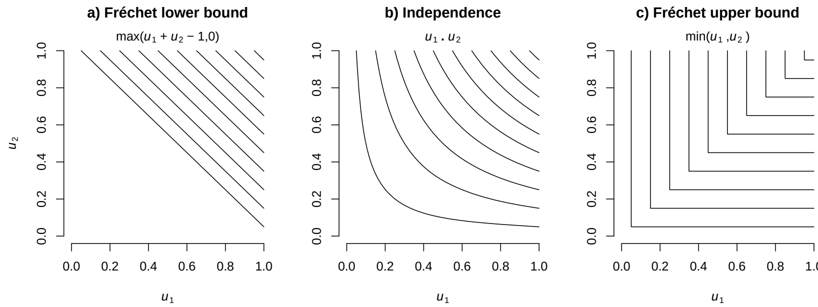

Fig. 4.3. Contour plots of bivariate cumulative probabilities: a) the Fréchet lower bound, b) independence, and c) the Fréchet upper bound.

$$\lim \frac{\psi_\theta(s_1)}{\psi'_\theta(s_2)} = \frac{\psi(s_1)}{\psi'(s_2)}$$

which denotes the appropriate one-sided limit as θ approaches an endpoint of the parameter interval Θ. □

(2) THEOREM 4.7 *[Nelsen, 1999, p. 113 (proof)]*
Let $C_{\mathrm{M}}(\mathbf{u};\theta)$ for $\theta \in \Theta$ be a family of Archimedean (survival) copulas with differentiable generators $\psi_\theta(s)$ in $\boldsymbol{\Omega}$. Then the limit of $C_{\mathrm{M}}(\mathbf{u};\theta)$ is equal to the F_U if and only if for s in $(0,1)$,

$$\lim \frac{\psi_\theta(s)}{\psi'_\theta(s)} = 0$$

which denotes the appropriate one-sided limit as θ approaches an endpoint of the parameter interval Θ. □

We then introduce the concept of bivariate tail dependence which relates to the amount of dependence in the upper or lower quadrant tail of a bivariate distribution. This concept is relevant to dependence in extreme values (which depends mainly on the tails). The definition will be given in terms of copulas due to invariance to increasing functions.

DEFINITION 4.8 *[Joe, 1997, p. 33]*
If a bivariate copula C_{B} is such that

$$\lim_{u\to 1} \frac{1-2u+C_{\mathrm{B}}(u,u)}{1-u} = \lambda_U$$

exists then C_{B} has upper tail dependence if $\lambda_U \in (0,1]$ and none if $\lambda_U = 0$. Similarly, if

$$\lim_{u\to 1} \frac{C_{\mathrm{B}}(u,u)}{u} = \lambda_L$$

exists then C_{B} has lower tail dependence if $\lambda_L \in (0,1]$ and none if $\lambda_L = 0$. □

Next, we consider two very specialized but interesting properties.

(1) THEOREM 4.9 *[Frank, 1979 (proof); Genest and MacKay, 1986a; Joe, 1997, pp. 140 and 143; Nelsen, 1999, pp. 31–33 and 107]*

Radial or reflection symmetry is obtained if and only if

$$C_{\mathrm{B}}(u_1,u_2) = \widehat{C}_{\mathrm{B}}(u_1,u_2)$$

where the survival copula $\widehat{C}_{\mathrm{B}}(u_1,u_2)$ has cdf margins. Hence, the copula $C_{\mathrm{B}}(u_1,u_2)$ or survival copula $\widehat{C}_{\mathrm{B}}(u_1,u_2)$ are symmetric with respect to the line $y = 1-x$. □

The Frank family (n°3 in Appendix C.1) is the only bivariate Archimedean (survival) copula to have this property. Note that this property does not hold for its permutation symmetric multivariate extension.

(2) DEFINITION 4.10 *[Nelsen, 1999, pp. 12 and 96]*

A family of (survival) copulas which includes the F_L, the P, and the F_U is called comprehensive. □

Now, we move on to defining max-infinite and min-infinite divisibility. This property will be most important in Subsection 4.3.5 to introduce higher order dependencies.

DEFINITION 4.11 *[Joe, 1997, p. 30]*

For a univariate cdf F, all positive powers F^{κ} for $\kappa > 0$ are cdfs. This is not always the case for multivariate cdfs. In general, for a multivariate cdf F_{M}, F_{M}^{κ} is a multivariate cdf for all $\kappa \geq m-1$. If F_{M}^{κ} is a multivariate cdf for all $\kappa > 0$ then F_{M} is max-infinitely divisible. □

The equivalent relationships are also valid for a survival function S and a multivariate survival function S_{M}. However, if S_{M}^{κ} is a multivariate survival function for all $\kappa > 0$ then S_{M} is min-infinitely divisible (Joe, 1997, p. 30).

THEOREM 4.12 *[Joe, 1997, pp. 9 and 101 (proof)]*

If we define $C_{\mathrm{M}}(\mathbf{u})$ to be an Archimedean (survival) copula with Laplace transform $\phi(s)$ then

- $C_{\mathrm{M}}^{\kappa}(\mathbf{u})$ is max-infinitely divisible up to dimension m if $-\ln[\phi(s)] \in L_m$ and
- $C_{\mathrm{M}}^{\kappa}(\mathbf{u})$ is max-infinitely divisible for all m if $-\ln[\phi(s)] \in L_\infty$

where L_n is the class of functions ω mapping $[0,\infty)$ to $[0,\infty)$ given that

- $\omega(0) = 0$,
- $\omega(\infty) = \infty$,
- and the n partial derivatives of ω have alternating signs

for $n = 1, 2\ldots,\infty$.

Thus, this property holds for Archimedean (survival) copulas obtained with a strict generator. □

Finally, we define concordance ordering respectively for the bivariate and the multivariate cases.

- DEFINITION 4.13 *[Joe, 1997, p. 36]*
  Let F, $\widetilde{F} \in F\,(F_1, F_2)$ where F_1 and F_2 are univariate cdfs. Then $\widetilde{F}$ is more concordant than F, if

  $$F(x_1, x_2) \leq \widetilde{F}(x_1, x_2) \qquad \forall\; x_1, x_2 \in (-\infty, \infty)$$

  The equivalent relationship exists for survival functions by replacing the cdfs with survival functions. □

- In the multivariate case ($m \geq 3$), the orderings of cdfs and survival functions are no longer equivalent. Therefore, there are various possible versions that could be considered as a multivariate dependence ordering.
  DEFINITION 4.14 *[Joe, 1997, pp. 37 and 39]*
  Let F_{M}, $\widetilde{F}_{\mathrm{M}} \in F\,(F_1, \ldots, F_m)$ where $F_1, \ldots, F_m$ are univariate cdfs. Then $\widetilde{F}_{\mathrm{M}}$ is more concordant than F_{M}, if

  $$F_{\mathrm{M}}(\mathbf{x}) \leq \widetilde{F}_{\mathrm{M}}(\mathbf{x}) \qquad \forall\; \mathbf{x} \in \mathbf{R}^m$$

  and

  $$S_{\mathrm{M}}(\mathbf{x}) \leq \widetilde{S}_{\mathrm{M}}(\mathbf{x}) \qquad \forall\; \mathbf{x} \in \mathbf{R}^m$$

  both hold.
  Then bivariate concordance, when $\widetilde{F}_{\mathrm{M}}$ is more concordant than F_{M}, implies that for all $1 \leq i < j \leq m$

  $$F_{\mathrm{B}}(x_i, x_j) \leq \widetilde{F}_{\mathrm{B}}(x_i, x_j) \qquad \forall\; x_i, x_j$$

  where F_{B} and $\widetilde{F}_{\mathrm{B}}$ are the (i, j) bivariate margins. □

Note that by the term "concordance", we mean that if $\widetilde{\mathbf{X}} \sim \widetilde{F}$ and $\mathbf{X} \sim F$ then the components of $\widetilde{\mathbf{X}}$ are more likely than those of $\mathbf{X}$ to take small values (or large values) simultaneously.

Further properties for bivariate and multivariate dependence ordering can be found in Joe (1997, pp. 38–39).

4.3.2 INSERTING MATCHING MIXTURE MARGINS

Before trying to extend the number of dependence parameters available, a method to obtain simple forms of multivariate distributions from Archimedean (survival) copulas will be considered. From the definition in Subsection 2.2.2, a (survival) copula is a function describing the dependence structure among $U(0, 1)$ margins. This implies that the margins must be chosen in a way such that they have their range on $[0, 1]$. Usually, univariate cdfs or survival functions are chosen.

In Section 2.1, Laplace transforms were found to be univariate mixture cdfs or univariate mixture survival functions. In the univariate context, mixtures are often used to introduce an additional parameter in order to model overdispersion.

Hence when mixtures are inserted as margins of Archimedean (survival) copulas, the multivariate distributions obtained have as general form

$$C_{\mathrm{M}}\left(\xi_1\left(y_1\right), \ldots, \xi_m\left(y_m\right)\right) = \phi\left(\sum_{j=1}^{m} \phi^{-1}\left(\xi_j\left(u_j\right)\right)\right)$$

where ξ_j and ϕ are Laplace transforms respectively corresponding to the univariate mixture margin j and the Archimedean (survival) copula.

If these Laplace transforms are all chosen to be the same with identical parameters, these Archimedean (survival) copulas reduce to the simpler general form

$$C_{\mathrm{M}}\left(\xi_1\left(y_1\right), \ldots, \xi_m\left(y_m\right)\right) = \phi\left(\sum_{j=1}^{m} y_j\right)$$

which at independence yields $\prod_{j=1}^{m} \phi(u_j)$ a product of mixtures. This also works with transformed univariate mixture margins, such as $\xi_j\left(-\ln[u_j]\right)$ or $\xi_j\left(H\left(y_j\right)\right)$, which have general form

$$C_{\mathrm{M}}\left(\xi_1\left(-\ln\left[F\left(y_1\right)\right]\right), \ldots, \xi_m\left(-\ln\left[F\left(y_m\right)\right]\right)\right) = \phi\left(-\sum_{j=1}^{m} \ln[u_j]\right)$$

$$C_{\mathrm{M}}\left(\xi_1\left(H\left(y_1\right)\right), \ldots, \xi_m\left(H\left(y_m\right)\right)\right) = \phi\left(\sum_{j=1}^{m} H(y_j)\right)$$

These Archimedean (survival) copulas or multivariate distributions can also be interpreted as multivariate Laplace transforms, see Equations (2.24) to (2.26) in Subsection 2.2.3.

4.3.3 EXTENSION TO TWO-PARAMETER FAMILIES

Two-parameter families of Archimedean (survival) copulas can be generated in a similar way to the one-parameter case. These can be obtained using generators (Nelsen, 1999, pp. 115–116) given by

$$\psi_{\alpha,\beta}(t) = \psi_{1,\beta} \circ \psi_{\alpha,1}(t) = \psi_{1,\beta}\left(\psi_{\alpha,1}(t)\right) = \left[\psi(t^{\alpha})\right]^{\beta} \tag{4.3}$$

which are composites of Equations (4.2) where $\circ$ represents a composition (Salas and Hille, 1990, p. 32). An example is listed in Appendix C.2 (n$^\circ$1).

Another method is to combine two Laplace transforms together to obtain a two-parameter inverse generator.

THEOREM 4.15 *[Joe, 1997, p. 374 (proof)]*

If $\psi_\delta(s)$ is a Laplace transform such that $-\ln[\psi_\delta(s)] \in L_\infty$ and $\phi_\theta(s)$ is another Laplace transform then

$$\eta_{\delta\theta}(s) = \phi_\theta\left(-\ln[\psi_\delta(s)]\right) \tag{4.4}$$

is a Laplace transform. □

Seven examples of two-parameter families of Archimedean (survival) copulas generated this way are shown in Appendix C.2.

Of course this method could be used, to some extent, recursively to build up different types of dependencies.

Finally, two-parameter families of Archimedean (survival) copulas can also be obtained from a simplification of the higher order dependence (survival) copulas presented in Subsections 4.3.4 and 4.3.5.

4.3.4 MIXTURES OF ARCHIMEDEAN COPULAS

More general families of Archimedean (survival) copulas can be generated by mixtures of powers. This method has already been applied in Equations (2.27) to generate Archimedean (survival) copulas. Because powers of Archimedean (survival) copulas are multivariate (survival functions) cdfs (see Subsection 4.3.1), these can in turn be used as part of mixtures of powers.

DEFINITION 4.16 *[Joe, 1997, pp. 87–89 and 155–157]*

A possible trivariate generalization of bivariate Archimedean (survival) copulas is

$$\begin{aligned}F_{\mathrm{T}}(-\ln[\zeta_\beta(-\ln[\widetilde{u}_1],-\ln[\widetilde{u}_2])],-\ln[\widetilde{u}_3]) &= \int_0^\infty \zeta^\beta(-\ln[\widetilde{u}_1],-\ln[\widetilde{u}_2])\,\widetilde{u}_3^\beta \,\mathrm{d}F_1(\beta)\\ &= \int_0^\infty\int_0^\infty \widetilde{u}_1^\alpha\widetilde{u}_2^\alpha \,\mathrm{d}F_2(\alpha;\beta)\widetilde{u}_3^\beta\,\mathrm{d}F_1(\beta)\end{aligned}$$

where:
- $F_2(\alpha;\beta)$ is the distribution with Laplace transform $\zeta_\beta(\alpha)$
- $\zeta_\beta(s) = \mathrm{e}^{-\beta\psi_\delta^{-1}(\phi_\theta(s))} = \mathrm{e}^{-\beta\left[\psi_\delta^{-1}\circ\phi_\theta(s)\right]} = \zeta^\beta(s)$

Then by recursively applying Equations (2.27), the corresponding trivariate (survival) copula is

$$C_{\mathrm{T}}(u_1,u_2,u_3;\delta,\theta) = \psi_\delta\left(\psi_\delta^{-1}\circ\phi_\theta\left(\phi_\theta^{-1}(u_1)+\phi_\theta^{-1}(u_2)\right)+\psi_\delta^{-1}(u_3)\right)$$

with intermediate step

$$\psi_\delta\left(-\ln[\zeta(-\ln[\widetilde{u}_1],-\ln[\widetilde{u}_2])]+\psi_\delta^{-1}(u_3)\right)$$

where the composition $\psi^{-1}\circ\phi\in L_\infty$ in order for $C_{\mathrm{T}}(u_1,u_2,u_3;\delta,\theta)$ to be a trivariate distribution. □

Notice that this method only produces (survival) copulas with at most $(m-1)$ distinct dependencies which are no longer symmetric. But on the other hand, these mixtures of Archimedean (survival) copulas have several desirable properties.

- PROPERTY 4.17 *[Joe, 1997, p. 87]*
  All the bivariate margins of mixtures of Archimedean (survival) copulas have the form of a bivariate Archimedean (survival) copula. ○
- PROPERTY 4.18 *[Joe, 1997, pp. 88 and 163]*
  Mixtures of Archimedean (survival) copulas simplify to Archimedean (survival) copulas when the same Laplace transform with the same dependence

parameter is used for both the mixture and the Archimedean (survival) copula. ○

- COROLLARY 4.19 *[Joe, 1997, pp. 90 and 156 (proof)]*
  Let $C_{\mathrm{M_i}}(\mathbf{u}) = \phi_i\left(\sum_{j=1}^{m} \phi_i^{-1}(u_j)\right)$ where ϕ_i is a Laplace transform with dependence parameter θ_i for $i = 1,2$. Suppose $\phi_1^{-1} \circ \phi_2 \in L_\infty$ then C_{M_2} is more concordant than C_{M_1}. ○
- COROLLARY 4.20 *[Joe, 1997, pp. 101–102 (proof)]*
  If we define $C_{\mathrm{M}}(\mathbf{u})$ to be a partially symmetric (survival) copula with compositions $\psi_i^{-1} \circ \phi_i \in L_{n_i}$ for sufficiently large n_i (greater than or equal to the number of terms in the argument of composition i) then $C_{\mathrm{M}}^{\kappa}(\mathbf{u})$ is max-infinitely divisible. ○

Generalizations to higher dimensions can be obtained if at each level of nesting of Laplace transforms (say ϕ_r within ϕ_s) the condition $\psi_s^{-1} \circ \phi_r \in L_\infty$ is satisfied in order for the result to be a multivariate distribution.

DEFINITION 4.21 *[Joe, 1997, pp. 87–89 and 155–157]*
In four dimensions, there are two possible nestings of Laplace transforms.

$$
\begin{aligned}
&C_{\mathrm{F}}(u_1,u_2,u_3,u_4;\theta_{12},\delta,\nu) = \\
&\quad \psi_\nu\left(\psi_\nu^{-1} \circ \phi_\delta\left(\phi_\delta^{-1} \circ \xi_{\theta_{12}}\left(\xi_{\theta_{12}}^{-1}(u_1) + \xi_{\theta_{12}}^{-1}(u_2)\right) + \phi_\delta^{-1}(u_3)\right) + \psi_\nu^{-1}(u_4)\right) \\
&C_{\mathrm{F}}(u_1,u_2,u_3,u_4;\theta_{12},\theta_{34},\theta) = \\
&\quad \psi_\theta\left(\psi_\theta^{-1} \circ \phi_{\theta_{12}}\left(\phi_{\theta_{12}}^{-1}(u_1) + \phi_{\theta_{12}}^{-1}(u_2)\right) + \psi_\theta^{-1} \circ \xi_{\theta_{34}}\left(\xi_{\theta_{34}}^{-1}(u_3) + \xi_{\theta_{34}}^{-1}(u_4)\right)\right) \quad (4.5)
\end{aligned}
$$

There is therefore no simple multivariate generalization, because in higher dimensions there are many possible nestings. □

Hence, this method of introducing additional parameters yields (survival) copulas with partial symmetry which are for instance suitable to model clustered data. Care must be taken because the clusters must be nested and cannot overlap (for an overlapping structure see Subsection 4.3.5).

Clustering The general form given by Equation (4.5) appears to be the most suitable for clustering.

As a first illustration, we shall consider the situation of a multicentre trial where patients have been recruited in several clinics and several measurements were collected during the same visit. Let there be o clinics denoted by k ($k = 1,\ldots,o$), n_k patients per clinic ($j = 1,\ldots,n_k$), and m_{jk} observations per patient ($i = 1,\ldots,m_{jk}$). Then ϕ_j and ψ must respectively be multivariate Laplace transforms of dimensions m_{jk} and n_k. This generates

$$
C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\theta}) = \psi\left(\sum_{j=1}^{n_k} \psi^{-1} \circ \phi_j\left(\sum_{i=1}^{m_{jk}} \phi_j^{-1}(u_{ijk})\right)\right) \quad (4.6)
$$

where ϕ_j and ψ respectively have dependence parameters θ_j and θ. These are contained in the dependence parameter vector $\boldsymbol{\theta}$ of the (survival) copula. Notice that for two clinics each containing two patients, Equation (4.5) is obtained.

Observations are exchangeable or permutable within a patient (u_{ijk} within ϕ_j), as are entire patients (ϕ_j within ψ). The parameter θ_j can be considered as measuring the dependence between patient j's observations or within patient j, whereas the parameter θ can be considered as measuring the between patient dependence (within clinics).

As a second illustration, consider the situation of a multicentre trial where several different measurements are collected per patient at each point in time. Let there be o clinics denoted by k ($k = 1, \ldots, o$), n_k patients per clinic ($j = 1, \ldots, n_k$), m_{jk} different variables collected per patient ($i = 1, \ldots, m_{jk}$), and l_{ijk} measurements over time per different variable collected for each patient ($h = 1, \ldots, l_{ijk}$). Then ξ_{ij}, ϕ_j, and ψ must respectively be multivariate Laplace transforms of dimensions l_{ijk}, m_{jk}, and n_k. This generates

$$C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\theta}) = \psi\left(\sum_{j=1}^{n_k} \psi^{-1} \circ \phi_j \left(\sum_{i=1}^{m_{jk}} \phi_j^{-1} \circ \xi_{ij} \left(\sum_{h=1}^{l_{ijk}} \xi_{ij}^{-1}\left(u_{hijk}\right)\right)\right)\right) \quad (4.7)$$

where ξ_{ij}, ϕ_j, and ψ respectively have dependence parameters θ_{ij}, θ_j, and θ. These are contained in the dependence parameter vector $\boldsymbol{\theta}$ of the (survival) copula.

Measurements over time within each variable collected per patient are exchangeable or permutable (u_{hijk} within ξ_{ij}, which is probably not a reasonable assumption because of changes over time), so is each variable collected per patient (ξ_{ij} within ϕ_j), and entire patients (ϕ_j within ψ). The parameter θ_{ij} can be considered as measuring the dependence between measurements over time of variable i observed on patient j. The parameter θ_j can be considered as measuring the between variable dependence for patient j. And parameter θ can be considered as measuring the between patient dependence.

Finally, it must be kept in mind that all Laplace transforms can be chosen to be the same. This particular choice will only simplify to an Archimedean (survival) copula if the dependence parameters at all levels become equal. The model will then become permutation symmetric in all its arguments. Hence, such (survival) copulas are suitable models for multi-level clustered data.

4.3.5 MIXTURES OF POWERS OF MAX-INFINITE DIVISIBLE DISTRIBUTIONS

Finally, families of (survival) copulas can be obtained with a dependence parameter for each bivariate margin as well as higher order dependencies. This is obtained by further extending the definition of mixtures of powers to min-infinite and max-infinite divisible distributions (Joe, 1997, pp. 30–31 and 98–101).

This extension is straightforward from the definition of a min-infinite or max-infinite divisible distribution introduced in Subsection 4.3.1. Hence from Equations (2.27), we can obtain

$$S_{\mathrm{M}}(\mathbf{v}) = \int_0^\infty \prod_{j=1}^{n} L_j^\lambda(\mathbf{v}) \mathrm{d}F(\lambda) = \phi\left(-\sum_{j=1}^{n} \ln[L_j(\mathbf{v})]\right)$$
$$F_{\mathrm{M}}(\mathbf{u}) = \int_0^\infty \prod_{j=1}^{n} U_j^\lambda(\mathbf{u}) \mathrm{d}F(\lambda) = \phi\left(-\sum_{j=1}^{n} \ln[U_j(\mathbf{u})]\right) \tag{4.8}$$

where $L_j(\mathbf{v})$ and $U_j(\mathbf{u})$ are respectively min-infinite and max-infinite divisible distributions for $j = 1, \ldots, n$ (with $n \leq m$).

Because Archimedean (survival) copulas and mixtures of these are all max-infinite (min-infinite) divisible distributions, this can be used to introduce an additional global dependence through the parameter of the Laplace transform ϕ.

DEFINITION 4.22 *[Joe, 1997, pp. 98–99]*

Let z_i be $U_i(u_i)$ or $L_i(v_i)$ and let $K_{ij}(z_i, z_j)$ be bivariate (survival) copulas that are (min-infinite) max-infinite divisible for $1 \leq i < j \leq m$.
Then consider the mixture

$$\int_0^\infty \prod_{i<j} K_{ij}^\lambda(z_i, z_j) \prod_{i=1}^{m} z_i^{\nu_i \lambda} \mathrm{d}F(\lambda) = \phi\left(-\sum_{i<j} \ln[K_{ij}(z_i, z_j)] - \sum_{i=1}^{m} \nu_i \ln[z_i]\right)$$

which has univariate margins of the form $F_i = \phi\left(-\frac{\ln[z_i]}{p_i}\right)$ where $p_i = \frac{1}{\nu_i + m - 1}$.
This mixture will be a (survival) copula if z_i is chosen to be $\mathrm{e}^{-p_i \phi^{-1}(u_i)}$ for $i = 1, \ldots, m$. Inserting these yields the (survival) copula

$$C_{\mathrm{M}}(\mathbf{u}; \boldsymbol{\delta}, \theta, \boldsymbol{\nu}) =$$
$$\phi_\theta\left(-\sum_{i<j} \ln\left[K_{ij}\left(\mathrm{e}^{-p_i \phi_\theta^{-1}(u_i)}, \mathrm{e}^{-p_j \phi_\theta^{-1}(u_j)}; \delta_{ij}\right)\right] + \sum_{i=1}^{m} \nu_i p_i \phi_\theta^{-1}(u_i)\right) \tag{4.9}$$

where the parameter θ of the Laplace transform ϕ can be seen as a global dependence measuring the minimal level or common dependence. The parameter δ_{ij} of each bivariate (survival) copula $K_{ij}(z_i, z_j)$ can be considered to model the pairwise dependence beyond this global dependence. These are contained in the dependence parameter vector $\boldsymbol{\delta}$ of the (survival) copula. On the other hand, the ν_i parameters contained in the vector $\boldsymbol{\nu}$ lead to bivariate and multivariate asymmetry. □

It is important to notice the difference with mixtures of Archimedean (survival) copulas introduced in Subsection 4.3.4 where the (survival) copulas could not overlap.

Hence from Definition 4.22, we are now able to generate families of (survival) copulas with $1 + \frac{m(m-1)}{2}$ dependencies and m asymmetry parameters. But two useful simplifications will still be considered.

(1) The asymmetries can be removed by setting all ν_i to zero.

$$C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\delta},\theta) = \phi_\theta\left(-\sum_{i<j}\ln\left[K_{ij}\left(\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_i)}{m-1}},\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_j)}{m-1}};\delta_{ij}\right)\right]\right) \quad (4.10)$$

(2) Another method of generating two-parameter families of (survival) copulas (with and without asymmetries) for Subsection 4.3.3 is to consider the case where this technique is applied to one Archimedean (survival) copula (Joe, 1997, pp. 103 and 149).

$$\begin{aligned} &C_{\mathrm{M}}(\mathbf{u};\delta,\theta,\boldsymbol{\nu}) = \\ &\quad \phi_\theta\left(-\ln\left[K\left(\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_1)}{\nu_1+1}},\ldots,\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_m)}{\nu_m+1}};\delta\right)\right] + \sum_{i=1}^{m}\frac{\nu_i\phi_\theta^{-1}(u_i)}{\nu_i+1}\right) \\ &C_{\mathrm{M}}(\mathbf{u};\delta,\theta) = \phi_\theta\left(-\ln\left[K\left(\mathrm{e}^{-\phi_\theta^{-1}(u_1)},\ldots,\mathrm{e}^{-\phi_\theta^{-1}(u_m)};\delta\right)\right]\right) \end{aligned} \quad (4.11)$$

It can be shown that the inverse generators for the (survival) copulas without asymmetries are obtained by combining the Laplace transform $\phi_\theta(s)$ with the one of the Archimedean (survival) copula K, say $\psi_\delta(s)$. This combining method was introduced in Equation (4.4) and produces two-parameter Laplace transforms of the form $\eta_{\delta\theta}(s) = \phi_\theta(-\ln[\psi_\delta(s)])$.
In this special case where no overlapping occurs, we then obtain (survival) copulas

$$C_{\mathrm{M}}(\mathbf{u};\delta,\theta) = \eta_{\delta\theta}\left(\sum_{j=1}^{m}\eta_{\delta\theta}^{-1}(u_j)\right)$$

which are also Archimedean (Joe, 1997, p. 150). Note that combining two identical Laplace transforms ϕ as well as keeping both parameters identical (θ) yields a one-parameter Laplace transform η_θ but does not reduce back to the Laplace transform ϕ_θ.

Mixtures of max-infinitely (min-infinite) divisible distributions are themselves max-infinitely (min-infinite) divisible.

COROLLARY 4.23 *[Joe, 1997, p. 102 (proof)]*

In general, consider

$$F_{\mathrm{M}}(\mathbf{u}) = \phi(-\ln[K(\mathbf{u})])$$

where $K(\mathbf{u})$ is a max-infinite (min-infinite) divisible distribution and $-\ln[\phi(s)] \in L_m$. Then $F_{\mathrm{M}}(\mathbf{u})$ is also max-infinitely (min-infinite) divisible. ○

Several interesting modifications can still be brought to Definition 4.22 either to obtain full dependencies or to obtain a specific type of dependence among the margins. Certain interesting possibilities will now be considered in the remaining subsection.

Clustering The general form presented in Equations (4.11) can be interpreted as a model with one level of nesting. To allow for several clusters, we obtain from Equations (4.8)

$$C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\theta}) = \phi_\theta\left(-\sum_{j=1}^{k}\ln\left[K_j\left(\mathrm{e}^{-\phi_\theta^{-1}(u_{1j})},\ldots,\mathrm{e}^{-\phi_\theta^{-1}\left(u_{m_j j}\right)};\theta_j\right)\right]\right) \quad (4.12)$$

where there are k clusters each containing m_j subjects.

Notice that these (survival) copulas are never Archimedean as all dependencies at the bivariate level are not available. Indeed, the different clusters are independent one from another. Therefore when the outer parameter (θ in this case) is set to independence, these (survival) copulas reduce to a function of a product of one-parameter Archimedean (survival) copulas.

Hence, this (survival) copula is different from the one of Equation (4.6) and can therefore be considered as an alternative way to model clustered data.

Higher levels of nesting can also be achieved as such (survival) copulas are themselves max-infinitely (min-infinite) divisible. This recursive method will be covered in the next subsection.

4.3.6 FULL UNSTRUCTURED DEPENDENCIES

The method just described in Subsection 4.3.5 can be generalized to combine (survival) copulas in an overlapping way. In the case where trivariate (survival) copulas K_{ijk} are combined without asymmetry parameters, this extends to

$$\begin{aligned} C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\delta},\boldsymbol{\epsilon},\theta) = \\ \phi_\theta\left(-\sum_{i<j<k}\ln\left[K_{ijk}\left(\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_i)}{C}},\mathrm{e}^{-\frac{\phi_\theta^{-1}\left(u_j\right)}{C}},\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_k)}{C}};\boldsymbol{\delta},\epsilon_{ijk}\right)\right]\right) \end{aligned} \quad (4.13)$$

where the general correction factor due to overlapping can be obtained by $C = \frac{(m-1)!}{(n-1)!(m-n)!}$ with n representing the size of the overlapping (survival) copula (for $n \leq m$). Thus in the present case, we have $C = \frac{m^2-3m+2}{2}$. The parameter vectors $\boldsymbol{\delta}$ and $\boldsymbol{\epsilon}$ respectively contain the bivariate and trivariate dependencies, and θ is a global dependence measuring the minimum level or common dependence.

It is now possible to use these mixtures of min-infinite or max-infinite divisible distributions recursively to obtain (survival) copulas with full unstructured dependencies. This produces (survival) copulas with $2^m - (m+1)$ dependence parameters.

Of course, the Laplace transforms involved can all be chosen to be identical in order to have all dependencies of the same type. Note that under no circumstances will these (survival) copulas collapse back to an Archimedean form.

For instance, a trivariate (survival) copula with all three bivariate dependencies and its trivariate dependency is obtained when the (survival) copula of Equation (4.9) has three margins. When the same Laplace transform is used, these have the form

$$
\begin{aligned}
C_{\mathrm{T}}\left(u_1,u_2,u_3;\delta_{12},\delta_{13},\delta_{23},\delta\right) = & \\
\phi_\delta \Bigg(-\ln\Bigg[\phi_{\delta_{12}}\Bigg(\phi_{\delta_{12}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_1)}{2}}\Bigg) + \phi_{\delta_{12}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_2)}{2}}\Bigg)\Bigg)\Bigg] & \\
-\ln\Bigg[\phi_{\delta_{13}}\Bigg(\phi_{\delta_{13}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_1)}{2}}\Bigg) + \phi_{\delta_{13}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_3)}{2}}\Bigg)\Bigg)\Bigg] & \\
-\ln\Bigg[\phi_{\delta_{23}}\Bigg(\phi_{\delta_{23}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_2)}{2}}\Bigg) + \phi_{\delta_{23}}^{-1}\Bigg(\mathrm{e}^{-\frac{\phi_\delta^{-1}(u_3)}{2}}\Bigg)\Bigg)\Bigg]\Bigg) & \quad (4.14)
\end{aligned}
$$

where δ is the trivariate dependence parameter.

A four dimension (survival) copula with all six bivariate dependencies, four trivariate dependencies, and its 4-variate dependency is obtained when the (survival) copula of Equation (4.13) has four margins and the trivariate (survival) copulas K_{ijk} have the form of the (survival) copula of Equation (4.9). When the same Laplace transform is used, these have as general form

$$
C_{\mathrm{F}}(\mathbf{u};\boldsymbol{\delta}) = \phi_\delta\left(-\sum_{i<j<k}\ln\left[\phi_{\delta_{ijk}}\left(-\sum_{l<o}\ln\left[\phi_{\delta_{lo}}\left(\phi_{\delta_{lo}}^{-1}(z_l) + \phi_{\delta_{lo}}^{-1}(z_o)\right)\right]\right)\right]\right)
$$

where the vector $\mathbf{u}$ contains four univariate marginal distributions, the vector $\boldsymbol{\delta}$ contains all dependence parameters including δ (the 4 way dependency), and z_l is of the form

$$
\mathrm{e}^{-\frac{1}{2}\phi_{\delta_{ijk}}^{-1}\left(\mathrm{e}^{-\frac{1}{3}\phi_\delta^{-1}(u_l)}\right)}
$$

for $l = i,j,k$.

Special parametrizations of (survival) copulas with full unstructured dependencies will now be discussed.

One dependency per level As a first simplification, we shall consider the case where all dependencies at each level are set to be equal. This results in a (survival) copula which has in total $(m-1)$ dependency parameters (one bivariate, one trivariate, ...).

The trivariate (survival) copula presented in Equation (4.14) becomes

$$
\begin{aligned}
C_{\mathrm{T}}(u_1,u_2,u_3;\delta,\theta) = & \\
\phi_\theta\left(-\sum_{i<j}\ln\left[\phi_\delta\left(\phi_\delta^{-1}\left(\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_i)}{2}}\right) + \phi_\delta^{-1}\left(\mathrm{e}^{-\frac{\phi_\theta^{-1}(u_j)}{2}}\right)\right)\right]\right) &
\end{aligned}
$$

with two parameters δ and θ respectively measuring a common bivariate dependence and the trivariate dependence.

Clustering Once again, models for clustered data can be generated. As we now have all dependencies at all levels available, multiple levels of nestings are obtained by setting the appropriate parameters equal. This can be achieved in two ways, with and without overlapping dependencies.

In the case of no overlapping, a (survival) copula with two levels of nesting, comparable to one presented in Equation (4.7), is obtained. This (survival) copula has as general form

$$C_{\mathrm{M}}(\mathbf{u};\boldsymbol{\theta}) = \phi\left(-\sum_{j=1}^{n_k} \ln\left[\phi_j\left(-\sum_{i=1}^{m_{jk}} \ln\left[K_{ij}\left(z_{1ijk},\ldots,z_{l_{ijk}ijk}\right)\right]\right)\right]\right)$$

where z_{hijk} is of the form

$$\mathrm{e}^{-\phi_j^{-1}\left(\mathrm{e}^{-\phi^{-1}\left(u_{hijk}\right)}\right)}$$

for $h = 1,\ldots,l_{ijk}$. The dependence parameter vector $\boldsymbol{\theta}$ contains θ, θ_j, and θ_{ij} respectively corresponding to the Laplace transforms ϕ, ϕ_j, and (survival) copulas K_{ij}. As in Equation (4.7), there are three levels of nesting among o clinics ($k = 1,\ldots,o$): n_k clusters corresponding to the number of patients per clinic ($j = 1,\ldots,n_k$), m_{jk} clusters corresponding to the number of different variables collected per patient ($i = 1,\ldots,m_{jk}$), and l_{ijk} clusters corresponding to the number of measurements over time per different variable collected for each patient ($h = 1,\ldots,l_{ijk}$).

Notice that when the parameter of the higher level of nesting (θ in this case) is set to independence, this (survival) copula does not reduce to Equation (4.12) although this corresponds to a logical two level nesting extension. Thus, the (survival) copulas presented in Equations (4.6) and (4.12) would be alternatives to consider if one of the two levels of nesting disappeared. On the other hand, the case of multiple levels of nesting with overlap can be obtained in many ways.

As an illustration, consider the situation of a continent wide multicentre trial. Then suppose the same dependence is expected among all different variables collected per patient but still allowing for within and between clinic dependencies. Hence, let there be o countries denoted by k ($k = 1,\ldots,o$), n_k clinics per country ($j = 1,\ldots,n_k$), m_{jk} patients per clinic ($i = 1,\ldots,m_{jk}$), and l different variables collected per patient. This requires the (survival) copula

$$C_{\mathrm{M}}(\mathbf{u};\delta,\boldsymbol{\theta}) = \phi\left(-\sum_{j=1}^{n_k} \ln\left[\phi_j\left(-\sum_{1\le r<s\le N} \ln\left[K_{rs}\left(z_r,z_s\right)\right]\right)\right]\right)$$

where N is the total number of observations $\left(l \times \sum_k \sum_j m_{jk}\right)$ and z_h is of the form

$$\mathrm{e}^{-\frac{\phi_j^{-1}(u_h)}{N-1}}$$

for $h = 1 \le r < s \le N$. The parameter δ captures the pairwise dependence among all different variables collected per patient regardless of the patient, clinic and country through the bivariate (survival) copula K_{rs}. On the other hand, the dependence parameter vector $\boldsymbol{\theta}$ contains θ_j and θ respectively measuring the dependence within and between clinics through the Laplace transforms ϕ_j and ϕ.

An alternative method would be to consider setting the appropriate parameters equal in (survival) copulas with full dependencies but this becomes rapidly unfeasible. Indeed, the dimension of a (survival) copula with full unstructured dependencies would have to correspond to the total number of observations (N). In other words, this (survival) copula would have N Laplace transforms where most of the dependence parameters would be set to independence.

Auto-regression and Markov chains This remaining type of dependence is useful in cases where there is a specific ordering, and perhaps a certain measure of distance, among the margins (such as for repeated measurements over time). The dependencies will then take some of the previous history of a subject into account as a decaying process (over time).

A first-order auto-regression dependence in continuous time consists in having only one bivariate dependency but where the dependence among margins is kept proportional to some measure of distance separating them. This measure of distance, say Δ, is usually chosen as $(t_j - t_i)$ the time difference between considered margins. This can be reduce to $(j - i)$ the difference between margin indices under a balanced design yielding a first-order stationary Markov chain. Hence, the strength of the bivariate dependencies is adjusted by setting the dependence parameter to ρ^{Δ}, $\mathrm{e}^{-\frac{\Delta}{\rho}}$, or $\mathrm{e}^{-\frac{\Delta^2}{\rho^2}}$ respectively called power, exponential and Gaussian structures.

Such a model could be obtained in several ways, but we shall just consider the case resulting from Equation (4.10). This yields a (survival) copula with general form

$$C_{\mathrm{M}}(\mathbf{u};\rho;t) = \phi\left(-\sum_{i<j} \ln\left[K_{ij}\left(\mathrm{e}^{-\frac{\phi^{-1}(u_i)}{m-1}}, \mathrm{e}^{-\frac{\phi^{-1}(u_j)}{m-1}}; \rho^{t_j - t_i}\right)\right]\right)$$

where the parameter of the Laplace transform ϕ is set to independence, the bivariate dependencies have a power structure, t_i is the observation time of margin i, and the time distance Δ is set to $(t_j - t_i)$ the time lag between margins j and i.

4.4 Further reading

5
Longitudinal models

This chapter will introduce multivariate models mainly for dependencies among time ordered observations, although in certain special cases the dependence no longer takes the ordering into account and is then also suitable for multiple responses or clustered data. Such models will be produced as products of conditional distributions, as described in Subsection 2.2.1, where each conditional distribution corresponds to a time point (or margin) given the previous one(s). Two different methods are considered to introduce dependencies among observations of the same subject.

The first approach consists in re-parameterizing the conditional distributions. This is performed by allowing some parameters to be a function of responses at the previous time points, creating sets of *updating equations*. The general approach is outlined in Section 5.1. In Section 5.2, the procedure for continuous variables is presented. The conditional distributions are all chosen to be different parametrizations of the same distribution yielding one general set of updating equations. Hence, different types of dependencies can be introduced through specific systems of updating equations. In Section 5.3, this is adapted for ordinal data.

The second methods allows the subjects to switch over time between a predefined number of different underlying or hidden states. All possible switches between hidden states at all available time points must then be considered. This is done by summing up the product of conditional distributions corresponding to each possible combination of states by time points together. A dependence among observations of the same subject is induced through a transition matrix containing the probabilities of changing from one hidden state to another. Note that the summation process just described expands exponentially with respect to the number of time points available and hidden states considered. Hence, the summation process must be ordered in a particular way to be computationally feasible; this is described in Section 5.4.

The models just described are fundamentally different from the marginal ones introduced in Chapter 4. In the case of marginal models such as (survival) copulas, all the information of a time point is collapsed together. On the other hand, the conditional models presented in this chapter are constructed dynamically, implying that the information is gathered (chronologically) throughout the series of each subject separately. This construction difference is very important because, in a marginal model, the regression constraints are applied to the underlying population (or "average") curve whereas in a conditional model they are applied to the individual curves. Note that these two types of regression constraints are identical in the case of continuous responses modeled with a linear regression with a Gaussian distribution but should not be assumed in general.

For certain dynamically updated models discussed in Section 5.2 and Subsection 5.3.2, the corresponding marginal model is found to be a known distribution, namely certain families of Archimedean (survival) copula from Section 4.3. Hence, the population probabilities of being in each category at each time point are available and can be obtained as described in Section 4.1. But additionally, for dynamic models, subject specific probabilities of being in each category at each time point can also be extracted. In the case of dynamic updates, Subsection 5.3.3 describes the method necessary to calculate individual or recursive probabilities for the two ordinal approaches. These probabilities are subject specific as they take into account the subject's previous history or variation captured by the set of updating equations. For hidden Markov chains, the end of Section 5.4 describes how to obtain the probability that a specific subject is in a particular hidden state given this subject's previous state history. The method to calculate the individual or recursive probabilities over time within each state is then described.

5.1 Dynamically updated models

Subsection 2.2.1 describes a method to generate multivariate densities from the product of independent univariate conditional distributions. Unfortunately, an extension to this approach which can be used to model nominal or ordinal responses is not straightforward.

Constructing multivariate distributions Extreme care must be taken in order to extend Equation (2.17) to products of cdfs. This can be shown by considering the bivariate case.

$$
\begin{aligned}
\frac{\partial^2 F_1(y_1)F_2(y_2|y_1)}{\partial y_1 \partial y_2} &= \frac{\partial^2 \int_{-\infty}^{y_1} f_1(y_1)\mathrm{d}y_1 \int_{-\infty}^{y_2} f_2(y_2|y_1)\mathrm{d}y_2}{\partial y_1 \partial y_2} \\
&= \frac{\partial \int_{-\infty}^{y_1} f_1(y_1)\mathrm{d}y_1 f_2(y_2|y_1)}{\partial y_1} \\
&= f_1(y_1)f_2(y_2|y_1) + \int_{-\infty}^{y_1} f_1(y_1)\mathrm{d}y_1 \frac{\partial f_2(y_2|y_1)}{\partial y_1} \\
&= f_\mathrm{B}(y_1,y_2) + F_1(y_1)\frac{\partial f_2(y_2|y_1)}{\partial y_1} \\
&\neq f_\mathrm{B}(y_1,y_2) \qquad (5.1) \\
&= \frac{\partial^2 F_\mathrm{B}(y_1,y_2)}{\partial y_1 \partial y_2}
\end{aligned}
$$

From this development, it is clear that $F_{\mathrm{B}}(y_1, y_2) \neq F_1(y_1)F_2(y_2|y_1)$ because the term $F_1(y_1)\frac{\partial f_2(y_2|y_1)}{\partial y_1}$ is generally different from zero unless y_1 and y_2 are independent. Nevertheless, this can be overcome under certain assumptions or solved for certain very particular cdfs.

This leads to the following two different alternatives:

(1) The multivariate distribution is assumed to be following an underlying continuous (or latent) process.
This hypothesis requires the conditional cdfs to be converted to conditional probabilities. Thus, it is necessary to adapt Equation (2.17) as

$$\begin{aligned}\Pr(Y_1 = y_1, \ldots, Y_m = y_m) = {} & \Pr(Y_1 = y_1)\Pr(Y_2 = y_2|Y_1 = y_1)\ldots \\ & \Pr(Y_m = y_m|Y_1 = y_1, \ldots, Y_{m-1} = y_{m-1})\end{aligned} \tag{5.2}$$

where $\Pr(Y_j = y_j|Y_1 = y_1, \ldots, Y_{j-1} = y_{j-1})$ is obtained by taking the appropriate univariate differences of the conditional cdfs corresponding to the categories of the ordinal data (the cumulative probability of the lower category is subtracted from the cumulative probability of the higher category). The multivariate likelihood is then obtained by multiplying together the result of Equation (5.2) for each subject.

(2) No assumptions are made about the discrete processes.
In this case, the full multivariate cdf is obtained. This is achieved by the product of certain conditional cdfs $F_j^*(y_j|y_1, \ldots, y_{j-1})$ which factor as $\frac{F_{\mathrm{M}}(y_1,\ldots,y_j)}{F_{\mathrm{M}}(y_1,\ldots,y_{j-1})}$. The problem illustrated in Equation (5.1) does not occur for these particular conditional cdfs as they do not correspond to the conditional densities of the multivariate distribution obtained.

$$\begin{aligned} f_j(y_j|y_1, \ldots, y_{j-1}) &= \tfrac{f_{\mathrm{M}}(y_1,\ldots,y_j)}{f_{\mathrm{M}}(y_1,\ldots,y_{j-1})} \\ F_j(y_j|y_1, \ldots, y_{j-1}) &\neq \tfrac{F_{\mathrm{M}}(y_1,\ldots,y_j)}{F_{\mathrm{M}}(y_1,\ldots,y_{j-1})} = F_j^*(y_j|y_1, \ldots, y_{j-1}) \end{aligned} \tag{5.3}$$

Hence, the multivariate likelihood is then obtained by taking the appropriate multivariate differences (see Section 4.1) for each subject and multiplying them together.

These cases will respectively be referred to as the latent and product of conditional cdfs approaches. But before proceeding on, several additional improvements are considered to allow these techniques to be more flexible and generate wider classes of multivariate distributions.

Generalizations First of all, this entire approach could alternatively be based on survival distributions due to their tractable relationships with cdfs. For univariate distributions, this was given in Equation (2.6) and extends to

$$f_{\mathrm{M}}(y_1, \ldots, y_m) = \frac{\partial^m F_{\mathrm{M}}(y_1, \ldots, y_m)}{\partial y_1 \ldots \partial y_m} = (-1)^m \frac{\partial^m S_{\mathrm{M}}(y_1, \ldots, y_m)}{\partial y_1 \ldots \partial y_m}$$

in the multivariate case.

Next, the conditional distributions can be chosen to be mixtures. These are of major interest as many can be obtained in closed form (see Chapter 3). More specifically, the gamma Laplace transform (or corresponding density) is chosen for illustrative purposes below (see Subsections 3.2.2 and 3.3.3). The choice to focus on mixtures is also motivated by the introduction of additional regression parameters. These are then used to induce dependence on the previous responses.

Until now, each conditional mixture has been assumed to be a function of the response y. Because this response is in the present case either nominal or ordinal, the method would be much more flexible if these conditional mixtures where given in terms of an arbitrary cdf. This requires the argument y of each of the conditional mixtures to be transformed to an arbitrary function $G(y)$ involving a cdf (or a survival function). However, care must be taken because this function $G(y)$ must be a strictly increasing transformation of y (in order not to affect the range of the conditional mixtures). For instance, the argument y could be replaced by an arbitrary integrated hazard $\widetilde{H}(y) = -\ln[\widetilde{S}(y)] = -\ln[1-\widetilde{F}(y)]$. Equations (2.8) and (2.9) illustrate such transformations applied to the argument of densities. As dependence on previous responses is only induced through the additional parameters of the mixture distributions, this generalization has the advantage that the conditional dependencies and the choice of the arbitrary function $G(y)$ remain independent.

Finally, recall from Subsection 2.2.1 that the conditional mixtures are independent one from another. This implies that each one could be chosen as a different distribution than the others. However only the case where they are all obtained from the same distribution will be considered here.

Hence, these changes allow families of multivariate distribution to be produced where the family is determined by the constructed joint function. In the case where joint cdfs (or survival functions) are constructed, this might seem to be another method of obtaining (survival) copulas (introduced in Subsection 2.2.2). Indeed these joint cdfs (or survival functions) are in terms of an arbitrary function $G(y)$ involving a univariate cdf (or survival function). But this will in general not be a (survival) copula unless the univariate cdfs (or survival functions) inserted also correspond to the marginal distributions of the resulting multivariate distribution.

5.2 Updating for continuous variables

The conditional mixture densities are directly used in Equation (2.17) to construct the multivariate distributions. Note that the arbitrary function $G(y)$ inserted into the mixture will be denoted by z and recall that the conditional densities are all chosen to be the same mixture density obtained from Equation (2.1).

DEFINITION 5.1 *[Bayes, 1958]*
Let us consider Bayes's formula

$$p(\lambda|z) = \frac{f(z|\lambda)p(\lambda)}{\int_0^\infty f(z|\lambda)p(\lambda)d\lambda} = \frac{f(z|\lambda)p(\lambda)}{f_{\mathrm{m}}(z)} \tag{5.4}$$

where $p(\lambda|z)$ is the posterior density, $f(z|\lambda)$ is a conditional density, $p(\lambda)$ is a prior density, and $f_{\mathrm{m}}(z)$ is a mixture density. □

If the subject's response history is taken into account, Equation (5.4) can be rewritten as

$$p(\lambda_{ij}|H_{ij}) = \frac{f(z_{ij}|\lambda_{ij}, H_{i,j-1})p(\lambda_{ij}|H_{i,j-1})}{f_{\mathrm{m}}(z_{ij}|H_{i,j-1})} \tag{5.5}$$

where H_{ij} denotes subject i's response history up to and including response j.

This sequential or recursive procedure yields a new distribution at each step by *updating*. This is particularly interesting when the posterior density has the same form as the prior density.

DEFINITION 5.2 *[Barnard, 1954; Wetherill, 1961]*
In Bayes's formula, if the posterior and prior densities have the same form, then the prior density is the conjugate of the conditional density. In this case, it is said to be *closed under sampling*. This also implies that the mixture must have a closed form. □

For closed under sampling cases (see Section 3.3), the posterior can be used as a prior for the next time point. The fact that the posterior remains of the same form as the prior implicitly suggests that the prior could be updated, by the subject's response history, to the posterior. This updating process is quite similar to autoregressive serial dependence where the mean at a specific time point is "corrected" or updated by the residual of the previous time point. These updating equations can then be deduced from the posterior density of Equation (5.5). Thus, the choice of the prior and therefore of the mixture is important in order to generate a sequence of densities of the same form.

Table 5.1 shows the updating equations for all the mixture densities closed under sampling presented in Section 3.3. These are obtained by comparing the mixing distribution (prior) to the re-parameterized distribution integrated out (posterior) in each of these mixtures.

Now, Equation (2.17) can be rewritten in terms of mixture densities.

$$f_{\mathrm{M}}(\mathbf{z}_i) = f_{\mathrm{m}}(z_{i1})f_{\mathrm{m}}(z_{i2}|H_{i1})\ldots f_{\mathrm{m}}(z_{im}|H_{i,m-1}) \tag{5.6}$$

Note that care must be taken with the first term of the series as it is unconditional and some of the parameters might not be identifiable.

Example In the case of an exponential density, Equation (B.5), with a two parameter gamma mixing density, Equation (B.8), or prior

Table 5.1. (Part 1) Updates for mixture densities.

Density	Mixing (prior)	Updates (from posterior)
$\mathrm{N}\left(\lambda,\psi^2\right)$	$\mathrm{N}\left(\kappa,\nu^2\right)$	$\kappa_{ij}=\frac{\psi^2\kappa_{i,j-1}+\nu^2_{i,j-1}z_{ij}}{\nu^2_{i,j-1}+\psi^2}$ $\nu^2_{ij}=\nu^2_{i,j-1}\frac{\psi^2}{\nu^2_{i,j-1}+\psi^2}$
$\mathrm{N}\left(\mu,\frac{1}{\lambda}\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+\frac{(z_{ij}-\mu)^2}{2}$ $\alpha_{ij}=\alpha_{i,j-1}+\frac{1}{2}$
$\mathrm{Po}\left(\lambda\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+1$ $\alpha_{ij}=\alpha_{i,j-1}+z_{ij}$
$\mathrm{Ga}\left(0,\frac{1}{\lambda},\alpha\right)$	$\mathrm{Ga}\left(0,\frac{1}{\phi},\psi\right)$	$\phi_{ij}=\phi_{i,j-1}+z_{ij}$ $\psi_{ij}=\psi_{i,j-1}+\alpha$
$\mathrm{E}\left(0,\frac{1}{\lambda}\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+z_{ij}$ $\alpha_{ij}=\alpha_{i,j-1}+1$
$\mathrm{S}\left(0,\frac{1}{\lambda},\kappa\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+G(z_{ij})^\kappa$ $\alpha_{ij}=\alpha_{i,j-1}+1$
$\mathrm{Gu}\left(\alpha+\beta\ln[\lambda\beta],\beta\right)$	$\mathrm{Ga}\left(0,\frac{1}{\phi},\psi\right)$	$\phi_{ij}=\phi_{i,j-1}+\beta\mathrm{e}^{-\frac{z_{ij}-\alpha}{\beta}}$ $\psi_{ij}=\psi_{i,j-1}+1$
$\mathrm{La}\left(\mu,\frac{1}{\lambda}\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+\lvert z_{ij}-\mu\rvert$ $\alpha_{ij}=\alpha_{i,j-1}+1$
$\mathrm{Si}\left(\mu,\frac{1}{\lambda}\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+\frac{(z_{ij}-\mu)^2}{2z_{ij}(1-z_{ij})\mu^2(1-\mu)^2}$ $\alpha_{ij}=\alpha_{i,j-1}+\frac{1}{2}$
$\mathrm{LN}\left(\mu,\frac{1}{\lambda},\theta\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+\frac{(\ln[z_{ij}-\theta]-\mu)^2}{2}$ $\alpha_{ij}=\alpha_{i,j-1}+\frac{1}{2}$

Here $\mathrm{N}(\text{mean},\text{variance})$ is a Gaussian distribution, $\mathrm{Po}(\text{mean = variance})$ is a Poisson distribution, $\mathrm{Ga}(\text{location},\text{scale},\text{shape})$ is a gamma distribution, $\mathrm{E}(\text{location},\text{scale})$ is an exponential distribution, $\mathrm{S}(\text{location},\text{parameter 1},\text{parameter 2})$ is a Weibull, Gompertz, or generalized Gompertz distribution where $G(y)$ is a function of y determining the distribution, $\mathrm{Gu}(\text{location},\text{scale})$ is a Gumbel distribution, $\mathrm{La}(\text{location},\text{scale})$ is a Laplace distribution, $\mathrm{Si}(\text{location},\text{scale})$ is a Simplex distribution, and $\mathrm{LN}(\text{location},\text{scale},\text{parameter 3})$ is a log-normal distribution.

$$p(\lambda_{ij}|H_{i,j-1})=\frac{\beta_{i,j-1}^{\alpha_{i,j-1}}\lambda_{ij}^{\alpha_{i,j-1}-1}\mathrm{e}^{-\beta_{i,j-1}\lambda_{ij}}}{\Gamma(\alpha_{i,j-1})}$$

the posterior density is also a two parameter gamma distribution (Subsection 3.3.3)

Table 5.1. (Part 2) Updates for mixture densities.

Density	Mixing (prior)	Updates (from posterior)
$\mathrm{Po}\left(\lambda\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2$ $\epsilon_{ij} = \epsilon_{i,j-1} + z_{ij}$
$\mathrm{Ga}\left(0, \frac{1}{\lambda}, \alpha\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2z_{ij}$ $\epsilon_{ij} = \epsilon_{i,j-1} + \alpha$
$\mathrm{E}\left(0, \frac{1}{\lambda}\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2z_{ij}$ $\epsilon_{ij} = \epsilon_{i,j-1} + 1$
$\mathrm{S}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2G(z_{ij})^{\kappa}$ $\epsilon_{ij} = \epsilon_{i,j-1} + 1$
$\mathrm{Gu}\left(\alpha + \beta \ln[\lambda\beta], \beta\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2\beta \mathrm{e}^{-\frac{z_{ij}-\alpha}{\beta}}$ $\epsilon_{ij} = \epsilon_{i,j-1} + 1$
$\mathrm{La}\left(\mu, \frac{1}{\lambda}\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2\lvert z_{ij} - \mu\rvert$ $\epsilon_{ij} = \epsilon_{i,j-1} + 1$

Here $\mathrm{GIG}(\text{location}, \text{scale}, \text{family})$ is a generalized inverse Gaussian distribution.

$$p(\lambda_{ij}|H_{ij}) = \frac{[\beta_{i,j-1} + z_{ij}]^{\alpha_{i,j-1}+1} \lambda_{ij}^{\alpha_{i,j-1}} \mathrm{e}^{-\lambda_{ij}[\beta_{i,j-1}+z_{ij}]}}{\Gamma(\alpha_{i,j-1}+1)} \tag{5.7}$$
$$= \frac{\beta_{ij}^{\alpha_{ij}} \lambda_{ij}^{\alpha_{ij}-1} \mathrm{e}^{-\beta_{ij}\lambda_{ij}}}{\Gamma(\alpha_{ij})}$$

and the mixture density is a Pareto distribution, see Equation (B.23) and Subsection 3.3.3,

$$f_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = \frac{\alpha_{i,j-1}\beta_{i,j-1}^{\alpha_{i,j-1}}}{[\beta_{i,j-1} + z_{ij}]^{\alpha_{i,j-1}+1}} \frac{\partial z_{ij}}{\partial y_{ij}} \tag{5.8}$$

where $\frac{\partial z_{ij}}{\partial y_{ij}}$ is the Jacobian.

PROPERTY 5.3

When z is chosen to be an arbitrary integrated hazard and α and β tend to infinity together, the Pareto distribution presented in Equation (5.8) reduces to the chosen corresponding density.

$$\lim_{\alpha=\beta\to\infty} \left[\frac{\alpha\beta^{\alpha}}{[\beta + \widetilde{H}(y)]^{\alpha+1}} \widetilde{h}(y) \right] = \mathrm{e}^{-\widetilde{H}(y)} \widetilde{h}(y) = \widetilde{f}(y)$$

This limit is equivalent to holding the location $\frac{\alpha}{\beta}$ of the gamma mixing distribution equal to unity while letting the variance $\frac{\alpha}{\beta^2}$ go to zero. ○

The updating equations can now be deduced from the posterior, Equations (5.7).

$$\begin{aligned}\beta_{ij} &= \beta_{i,j-1} + z_{ij} \\ \alpha_{ij} &= \alpha_{i,j-1} + 1\end{aligned} \tag{5.9}$$

Care must now be taken with $f_{\mathrm{m}}(z_{i1})$, the first term of Equation (5.6), because the two parameters β_{i0} and α_{i0} are not identifiable. Because the gamma mixing distribution has a multiplicative effect, its theoretical mean $\frac{\alpha}{\beta}$ should be held to unity. Hence in the initial state, these parameters are defined as $\beta_{i0} = \alpha_{i0} = \frac{1}{\delta}$.

From the updating equations, the sequence of conditional Pareto mixtures are

$$\begin{aligned}
f_{\mathrm{m}}(z_{i1};\delta) &= \frac{\frac{1}{\delta}\left(\frac{1}{\delta}\right)^{\frac{1}{\delta}}}{\left(\frac{1}{\delta}+z_{i1}\right)^{\frac{1}{\delta}+1}}\frac{\partial z_{i1}}{\partial y_{i1}} \\
f_{\mathrm{m}}(z_{i2}|H_{i1};\delta) &= \frac{\left(\frac{1}{\delta}+1\right)\left(\frac{1}{\delta}+z_{i1}\right)^{\frac{1}{\delta}+1}}{\left(\frac{1}{\delta}+z_{i1}+z_{i2}\right)^{\frac{1}{\delta}+2}}\frac{\partial z_{i2}}{\partial y_{i2}} \\
f_{\mathrm{m}}(z_{i3}|H_{i2};\delta) &= \frac{\left(\frac{1}{\delta}+2\right)\left(\frac{1}{\delta}+z_{i1}+z_{i2}\right)^{\frac{1}{\delta}+2}}{\left(\frac{1}{\delta}+z_{i1}+z_{i2}+z_{i3}\right)^{\frac{1}{\delta}+3}}\frac{\partial z_{i3}}{\partial y_{i3}} \\
&\vdots \\
f_{\mathrm{m}}(z_{im}|H_{i,m-1};\delta) &= \frac{\left(\frac{1}{\delta}+m-1\right)\left(\frac{1}{\delta}+\sum_{j=1}^{m-1}z_{ij}\right)^{\frac{1}{\delta}+m-1}}{\left(\frac{1}{\delta}+\sum_{j=1}^{m}z_{ij}\right)^{\frac{1}{\delta}+m}}\frac{\partial z_{im}}{\partial y_{im}}
\end{aligned}$$

Applying Equation (5.6), the multivariate Pareto density is

$$\begin{aligned}
f_{\mathrm{M}}(\mathbf{z}_i;\delta) &= \frac{\left(\frac{1}{\delta}\right)^{\frac{1}{\delta}}\frac{1}{\delta}\left(\frac{1}{\delta}+1\right)\ldots\left(\frac{1}{\delta}+m-1\right)}{\left(\frac{1}{\delta}+\sum_{j=1}^{m}z_{ij}\right)^{\frac{1}{\delta}+m}}\prod_{j=1}^{m}\frac{\partial z_{ij}}{\partial y_{ij}} \\
&= \frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)}\left(\frac{1}{\delta}\right)^{\frac{1}{\delta}}\left(\frac{1}{\delta}+\sum_{j=1}^{m}z_{ij}\right)^{-\frac{1}{\delta}-m}\prod_{j=1}^{m}\frac{\partial z_{ij}}{\partial y_{ij}}
\end{aligned} \tag{5.10}$$

which is invariant to reordering of the observations. It is therefore a frailty multivariate density, which is suitable for modelling clustered or multiple response data, and will from now on be referred to as the frailty multivariate Pareto density.

Although the updating equations obtained by applying Bayes's formula (see Table 5.1 for the mixtures closed under sampling in Section 3.3) appear to require an ordered sequence of observations, they all generate multivariate distributions which are invariant to reordering. These updating equations will from now on be referred to as frailty updates.

It can be seen from Table 5.1 (Part 1) that the Gaussian mixture obtained in Subsection 3.3.1 has updating equations quite different from any other distribution. Indeed, other distributions have one update accumulating the subject's response history (corresponding to another than the location one) while the other update acts as a counting process. In the Gaussian case, the update accumulating the subject's response history corresponds to the location parameter while the second update corresponding to the scale parameter mixes the variability of the two Gaussian distributions.

This difference can be explained by the way the conditional density binds to the mixing distribution. For the Gaussian distribution, the mixing distribution acts additively whereas it has a multiplicative effect (Laplace transform "style" mixture) for other distributions — see Equation (2.11) for the exponential case.

Updating equations Although the set of updates provided by Bayes's formula appears to be the most logical choice, many other updates are possible. Recall that the conditional mixtures of Equation (5.6) are independent. Hence, any reparametrization of these conditional mixtures can be applied to introduce dependence on previous responses. But in order to keep the same form for all conditional mixtures, the general structure of the updating equations provided by Bayes's formula must be kept.

For instance, the updates could be adapted in order to take the amount of elapsed time between two observations or a measurement of distance between two margins. Thus, a serial update can be generated by modifying the Pareto updates presented in Equations (5.9) as

$$\begin{aligned}\beta_{ij} &= \rho^{t_{ij}-t_{i,j-1}}\beta_{i,j-1} + \frac{(1-\rho^{t_{ij}-t_{i,j-1}})}{\delta} + z_{ij} \\ \alpha_{ij} &= \rho^{t_{ij}-t_{i,j-1}}\alpha_{i,j-1} + \frac{(1-\rho^{t_{ij}-t_{i,j-1}})}{\delta} + 1\end{aligned} \tag{5.11}$$

where ρ can be called the discount parameter as it reduces the influence of previous observations, $t_{ij} - t_{i,j-1}$ represents the elapsed time or a measurement of distance between observations $j-1$ and j for subject i, and δ is the initiation parameter (Lindsey, 1999, p. 72).

Similarly, a Markov update can be generated by

$$\begin{aligned}\beta_{ij} &= \frac{1}{\delta} + \rho^{t_{ij}-t_{i,j-1}}z_{i,j-1} + z_{ij} \\ \alpha_{ij} &= \rho^{t_{ij}-t_{i,j-1}}\alpha_{i,j-1} + \frac{(1-\rho^{t_{ij}-t_{i,j-1}})}{\delta} + 1\end{aligned} \tag{5.12}$$

In this case, the parameter β_{ij} depends only on its previous value discounted in function of the distance between the two observations, rather than an accumulation of all the subject's previous values (Lindsey, 1999, p. 72).

Independence is obtained for both of these dependencies by setting $\rho = 0$. Care must be taken when the dependence ρ gets large as this could indicate that there is little variability among the individual series. In this case, a uniform dependence over the succeeding observations should perhaps be considered instead. This is possible in the case of the serial update as the frailty dependence is obtained when $\rho \to 1$.

Linking conditional to marginal models The gamma Archimedean copula (Kimeldorf and Sampson, 1975) can be written as

$$\mathrm{C_M}(\mathbf{u};\delta) = \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}} \tag{5.13}$$

where u_j is a $U(0,1)$ cdf for margin j, $\mathbf{u}$ is a vector containing the univariate margins, m represents the number of univariate margins, and δ is the dependence between these margins (see Appendix C.1).

The corresponding multivariate density (Cook and Johnson, 1981) can then be obtained by taking partial derivatives with respect to each margins.

$$\begin{aligned}
\frac{\partial \mathrm{C_M}(\mathbf{u};\delta)}{\partial u_1} &= \left(\frac{\frac{1}{\delta}}{\frac{1}{\delta}}\right) \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-1} u_1^{-\delta-1} \\
\frac{\partial^2 \mathrm{C_M}(\mathbf{u};\delta)}{\partial u_1 \partial u_2} &= \left(\frac{\frac{1}{\delta}}{\frac{1}{\delta}}\right) \left(\frac{\frac{1}{\delta}+1}{\frac{1}{\delta}}\right) \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-2} (u_1 u_2)^{-\delta-1} \\
\frac{\partial^3 \mathrm{C_M}(\mathbf{u};\delta)}{\partial u_1 \partial u_2 \partial u_3} &= \left(\frac{\frac{1}{\delta}}{\frac{1}{\delta}}\right) \left(\frac{\frac{1}{\delta}+1}{\frac{1}{\delta}}\right) \left(\frac{\frac{1}{\delta}+2}{\frac{1}{\delta}}\right) \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-3} \prod_{j=1}^{3} u_j^{-\delta-1} \\
&\vdots \\
\frac{\partial^m \mathrm{C_M}(\mathbf{u};\delta)}{\partial u_1 \ldots \partial u_m} &= \frac{\frac{1}{\delta} \cdots \left(\frac{1}{\delta}+m-1\right)}{\left(\frac{1}{\delta}\right)^m} \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-m} \prod_{j=1}^{m} u_j^{-\delta-1} \\
&= \frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)\left(\frac{1}{\delta}\right)^m} \left[\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-m} \prod_{j=1}^{m} u_j^{-\delta-1} \qquad (5.14) \\
&= c_{\mathrm{M}}(\mathbf{u};\delta)
\end{aligned}$$

From Property 2.18 or Definition 2.19, the associated gamma Archimedean survival copula can be written as

$$\mathrm{C_M}(\mathbf{v};\delta) = \left[\sum_{j=1}^{m} v_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}} \tag{5.15}$$

where v_j is a survival function for margin j, and $\mathbf{v}$ is a vector containing the survival functions of the different univariate margins.

The corresponding multivariate density can then be obtained by taking partial derivatives with respect to each of the margins.

$$\begin{aligned} c_{\mathrm{M}}(\mathbf{v};\delta) &= \frac{\partial^m \mathrm{C_M}(\mathbf{v};\delta)}{\partial v_1 \dots \partial v_m} \frac{\partial[1-u_1]}{\partial u_1} \cdots \frac{\partial[1-u_m]}{\partial u_m} \\ &= (-1)^m \frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)\left(\frac{1}{\delta}\right)^m} \left[\sum_{j=1}^{m} v_j^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-m} \prod_{j=1}^{m} v_j^{-\delta-1} \end{aligned} \tag{5.16}$$

Now, recall from Subsection 4.3.2 that the Archimedean (survival) copula simplifies if the univariate margins are chosen to be mixtures obtained with the same Laplace transform.

In the particular case of the gamma Archimedean (survival) copula, these mixture margins are

$$\phi(z_j) = [1+\delta z_j]^{-\frac{1}{\delta}} \tag{5.17}$$

with Jacobians

$$-[1+\delta z_j]^{-\frac{1}{\delta}-1} \frac{\partial z_j}{\partial y_j} \tag{5.18}$$

(see Theorem 2.4 for the special case where z_j is an arbitrary integrated hazard). Hence, inserting these mixture margins into Equation (5.16) yields

$$\begin{aligned} f_{\mathrm{M}}(\mathbf{z}_i;\delta) &= (-1)^m \frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)\left(\frac{1}{\delta}\right)^m} \left[\sum_{j=1}^{m} \left[(1+\delta z_j)^{-\frac{1}{\delta}}\right]^{-\delta} - (m-1)\right]^{-\frac{1}{\delta}-m} \\ &\qquad \prod_{j=1}^{m} \left[(1+\delta z_j)^{-\frac{1}{\delta}}\right]^{-\delta-1} \prod_{j=1}^{m} \left[-(1+\delta z_j)^{-\frac{1}{\delta}-1} \frac{\partial z_j}{\partial y_j}\right] \\ &= \frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)} \left(\frac{1}{\delta}\right)^{\frac{1}{\delta}} \left(\frac{1}{\delta} + \sum_{j=1}^{m} z_{ij}\right)^{-\frac{1}{\delta}-m} \prod_{j=1}^{m} \frac{\partial z_{ij}}{\partial y_{ij}} \end{aligned}$$

which is identical to the frailty multivariate Pareto presented in Equation (5.10).

From Definition 4.4, the (survival) copula simplifies to the product of its margins at independence. In the case of the gamma Archimedean survival copula, this occurs as δ tends to zero. Hence at independence, the frailty multivariate Pareto density of Equation (5.10) yields a product of gamma Laplace transforms. In the special case where z_j is chosen to be an arbitrary integrated hazard, the frailty multivariate Pareto density reduces to the chosen corresponding density (see Property 5.3).

Now a similar relationship could be found for the gamma Archimedean copula given in Equation (5.14).

Recall from Property 2.18 and Definition 2.19 that Equations (5.13) and (5.15) are different multivariate distributions. Indeed, although their ranges are the same, their domains are inverted because cdfs go from zero to one and survival functions from one to zero. It can be seen by comparing multivariate densities given in Equations (5.14) and (5.15) that this domain inversion has led to an additional factor of $(-1)^m$.

Hence, a dynamic model corresponding to the gamma Archimedean copula can be constructed from conditional Pareto mixtures of the form

$$f_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = -\frac{\alpha_{i,j-1}\beta_{i,j-1}^{\alpha_{i,j-1}}}{[\beta_{i,j-1}+z_j]^{\alpha_{i,j-1}+1}}\frac{\partial z_{ij}}{\partial y_{ij}} \tag{5.19}$$

with frailty updates given in Equations (5.9) and where z_{ij} is an arbitrary strictly decreasing function (thus inverting the domain) such as $-\ln[F(y)]$. This will also have the form of a frailty multivariate Pareto density.

$$f_{\mathrm{M}}(\mathbf{z}_i;\delta) = (-1)^m\frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)}\left(\frac{1}{\delta}\right)^{\frac{1}{\delta}}\left(\frac{1}{\delta}+\sum_{j=1}^m z_{ij}\right)^{-\frac{1}{\delta}-m}\prod_{j=1}^m\frac{\partial z_{ij}}{\partial y_{ij}} \tag{5.20}$$

Hence, this multivariate density is obtained when mixture margins and Jacobian respectively given in Equations (5.17) and (5.18) are inserted into the gamma Archimedean copula (see Lemma 2.5 for the special case where z_j is minus the logarithm of an arbitrary cdf).

Next, Property 5.3 is adapted as follows.

PROPERTY 5.4

When z is chosen to be minus the logarithm of an arbitrary cdf and α and β tend to infinity together, the Pareto distribution presented in Equation (5.19) reduces to the chosen corresponding density.

$$\lim_{\alpha=\beta\to\infty}\left[-\frac{\alpha\beta^\alpha}{\{\beta-\ln[\widetilde{F}(y)]\}^{\alpha+1}}\left(-\frac{\widetilde{f}(y)}{\widetilde{F}(y)}\right)\right] = \mathrm{e}^{-\{-\ln[\widetilde{F}(y)]\}}\frac{\widetilde{f}(y)}{\widetilde{F}(y)} = \widetilde{f}(y)$$

As previously, this limit is equivalent to holding the location $\frac{\alpha}{\beta}$ of the gamma mixing distribution equal to unity while letting the variance $\frac{\alpha}{\beta^2}$ go to zero. ○

Hence at independence and in the special case where z_j is chosen to be minus the logarithm of an arbitrary cdf, the frailty multivariate Pareto density of Equation (5.20) also reduces to the chosen corresponding density.

Now, Cook and Johnson (1981) have related these two families of multivariate distributions to several well known cases.

- A multivariate generalized logistic distribution, Equation (B.53), is obtained by inserting gamma mixture cdf margins of the form

$$u_j = \left[1+\mathrm{e}^{-y_j}\right]^{-\frac{1}{\delta}}$$

with Jacobian

$$\frac{\mathrm{e}^{-y_j}}{\delta}\left[1+\mathrm{e}^{-y_j}\right]^{-\frac{1}{\delta}-1}$$

into the multivariate gamma copula, Equation (5.14).

- A multivariate Burr distribution, Equation (B.49), is obtained by inserting gamma mixture survival margins of the form

$$v_j = \left[1+\beta y_j^{\kappa}\right]^{-\frac{1}{\delta}}$$

with Jacobian

$$-\frac{\beta\kappa y_j^{\kappa-1}}{\delta}\left[1+\beta y_j^{\kappa}\right]^{-\frac{1}{\delta}-1}$$

into the multivariate survival gamma copula, Equation (5.16).

- A multivariate Pareto distribution, Equation (B.54), is obtained by inserting gamma mixture survival margins of the form

$$v_j = \left(\frac{y_j}{\theta}\right)^{-\frac{1}{\delta}}$$

with Jacobian

$$-\frac{1}{\delta\theta}\left(\frac{y_j}{\theta}\right)^{-\frac{1}{\delta}-1}$$

into the multivariate gamma survival copula, Equation (5.16).

- A multivariate Gaussian distribution

$$\frac{\Gamma\left(\frac{1}{\delta}+m\right)}{\Gamma\left(\frac{1}{\delta}\right)\left(\frac{1}{\delta}\right)^m}\left[\sum_{j=1}^{m}\Phi(y_j)^{-\delta}-(m-1)\right]^{-\frac{1}{\delta}-m}\prod_{j=1}^{m}\frac{f_G(y_j)}{\Phi(y_j)^{\delta+1}}$$

where $f_G(y)$ is the standardized Gaussian density and $\Phi(\cdot)$ is the corresponding cdf, respectively presented in Equations (B.42) and (B.43), is obtained by inserting Gaussian cdf margins of the form

$$u_j = \Phi(y_j)$$

with Jacobian

$$f_G(y_j)$$

into the multivariate gamma copula, Equation (5.14).

Similarly, all other frailty multivariate distributions obtained from the frailty updates of Table 5.1 can be linked to Archimedean (survival) copulas. In addition, these distributions conserve one of the main properties of (survival) copulas. Indeed at independence, the product of the univariate marginal distributions is obtained as both the dependence structure of the copula and of the mixture vanish.

5.3 Updating for ordinal responses

5.3.1 LATENT APPROACH

In this approach, the multivariate distribution is assumed to be following a latent process (Joe, 1997, pp. 221 and 236–237). Hence to build the likelihood, the appropriate differences of conditional mixture cdfs are used requiring Equation (2.17) to be adapted as shown in Equation (5.2). Note that the arbitrary function $G(y)$ inserted into the mixture will involve the cdf (or survival function) of one of the (nominal or ordinal) parametrizations of Section 1.3 and be denoted by z. Also recall that the conditional cdfs are all chosen to be the same mixture cdf obtained from Equation (2.2).

If the subject's response history is taken into account, the conditional mixture cdf can be written as

$$F_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = \int_0^\infty F(z_{ij}|\lambda_{ij}, H_{i,j-1})dF(\lambda_{ij}|H_{i,j-1})$$

where H_{ij} denotes subject i's response history up to and including response j. The conditional probabilities are then obtained by taking the appropriate differences

$$\Pr(Z_{ij} = k_j|H_{i,j-1}) = F_{\mathrm{m}}(Z_{ij} = k_j|H_{i,j-1}) - F_{\mathrm{m}}(Z_{ij} = k_j - 1|H_{i,j-1})$$

where $\Pr(Z_{ij} = k_j|H_{i,j-1})$ denotes the conditional probability of observation j from individual i being in category k for $k = 0, \ldots, (\mathtt{nc} - 1)$.

Now to be able to update recursively the conditional mixture cdf, we must ensure that the mixing distribution used matches the distribution integrated out to obtain a closed form.

DEFINITION 5.5

A closed form mixture cdf or mixture survival function is said to be *recursively closed*, if the distribution integrated out and the mixing distribution have the same form. □

Tables 5.2 and 5.3 respectively show the updating equations for the mixture cdfs (see Section 3.1) and mixture survival functions (see Section 3.2) that are recursively closed. These are obtained by comparing the mixing distribution to the re-parameterized distribution integrated out in each of these mixtures.

The joint density for individual i is then obtained by rewriting Equation (5.2) in terms of conditional probabilities.

Table 5.2. Updates for mixture cdfs.

CDF	Mixing	Updates
$\mathrm{Gu}\left(\alpha+\beta\ln[\lambda\beta],\beta\right)$	$\mathrm{E}\left(0,\frac{1}{\lambda}\right)$	$\phi_{ij}=\phi_{i,j-1}+\beta \mathrm{e}^{-\frac{z_{ij}-\alpha}{\beta}}$
$\mathrm{Gu}\left(\alpha+\beta\ln[\lambda\beta],\beta\right)$	$\mathrm{Ga}\left(0,\frac{1}{\beta},\alpha\right)$	$\beta_{ij}=\beta_{i,j-1}+\phi \mathrm{e}^{-\frac{z_{ij}-\psi}{\phi}}$
$\mathrm{Gu}\left(\alpha+\beta\ln[\lambda\beta],\beta\right)$	$\mathrm{IG}\left(1,\frac{1}{\phi}\right)$	$\phi_{ij}=\sqrt{\phi_{i,j-1}\left(\phi_{i,j-1}+2\beta \mathrm{e}^{-\frac{z_{ij}-\alpha}{\beta}}\right)}$
$\mathrm{Gu}\left(\alpha+\beta\ln[\lambda\beta],\beta\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}},\frac{1}{\phi},\epsilon\right)$	$\phi_{ij}=\phi_{i,j-1}+2\beta \mathrm{e}^{-\frac{z_{ij}-\alpha}{\beta}}$

Here Gu(location, scale) is a Gumbel distribution, E(location, scale) is an exponential distribution, Ga(location, scale, shape) is a gamma distribution, IG(location, scale) is an inverse Gaussian distribution, and GIG(location, scale, family) is a generalized inverse Gaussian distribution.

$$\Pr(Z_{i1}=k_1,\ldots,Z_{im}=k_m)= \\ \Pr(Z_{i1}=k_1)\Pr(Z_{i2}=k_2|H_{i1})\ldots\Pr(Z_{im}=k_m|H_{i,m-1}) \quad (5.21)$$

Notice that in this case, the conditional terms are also independent one from another. In this case, as previously care must be taken with the first term of the series as it is unconditional and some of the parameters might not be identifiable.

Example In the case of an exponential survival function, Equation (B.6), with a two parameter gamma mixing density, Equation (B.8), the distribution integrated out is also a two parameter gamma distribution (Subsection 3.2.2)

$$\begin{aligned} f(\lambda_{ij}|H_{ij}) &= \frac{[\beta_{i,j-1}+z_{ij}]^{\alpha}\lambda_{ij}^{\alpha-1}\mathrm{e}^{-\lambda_{ij}[\beta_{i,j-1}+z_{ij}]}}{\Gamma(\alpha)} \\ &= \frac{\beta_{ij}^{\alpha}\lambda_{ij}^{\alpha-1}\mathrm{e}^{-\beta_{ij}\lambda_{ij}}}{\Gamma(\alpha)} \end{aligned} \quad (5.22)$$

and the mixture survival function corresponds to a Pareto distribution, see Equation (B.24) and Subsection 3.2.2,

$$S_{\mathrm{m}}(z_{ij}|H_{i,j-1})=\left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}+z_{ij}}\right)^{\alpha}$$

The updating equation can now be deduced from the distribution integrated out, Equations (5.22).

$$\beta_{ij}=\beta_{i,j-1}+z_{ij}$$

Now, if the proportional odds parametrization of Subsection 1.3.5 is chosen, then the cumulative probability of being observed in category k is given by the cdf $F(Y_{ij}=k_j)=p_{ijk_j}$.

Table 5.3. Updates for mixture survival functions.

Survival function	Mixing	Updates
$\mathrm{E}\left(0, \frac{1}{\lambda}\right)$	$\mathrm{E}\left(0, \frac{1}{\beta}\right)$	$\beta_{ij} = \beta_{i,j-1} + z_{ij}$
$\mathrm{W}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{E}\left(0, \frac{1}{\beta}\right)$	$\beta_{ij} = \beta_{i,j-1} + z_{ij}^{\kappa}$
$\mathrm{Go}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}\right)$	$\mathrm{E}\left(0, \frac{1}{\beta}\right)$	$\beta_{ij} = \beta_{i,j-1} + \mathrm{e}^{\kappa z_{ij}}$
$\mathrm{GGo}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}, \nu\right)$	$\mathrm{E}\left(0, \frac{1}{\beta}\right)$	$\beta_{ij} = \beta_{i,j-1} + \mathrm{e}^{\kappa z_{ij}^{\nu}}$
$\mathrm{S}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{E}\left(0, \frac{1}{\beta}\right)$	$\beta_{ij} = \beta_{i,j-1} + G(z_{ij})^{\kappa}$
$\mathrm{E}\left(0, \frac{1}{\lambda}\right)$	$\mathrm{Ga}\left(0, \frac{1}{\beta}, \alpha\right)$	$\beta_{ij} = \beta_{i,j-1} + z_{ij}$
$\mathrm{W}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{Ga}\left(0, \frac{1}{\beta}, \alpha\right)$	$\beta_{ij} = \beta_{i,j-1} + z_{ij}^{\kappa}$
$\mathrm{Go}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}\right)$	$\mathrm{Ga}\left(0, \frac{1}{\beta}, \alpha\right)$	$\beta_{ij} = \beta_{i,j-1} + \mathrm{e}^{\kappa z_{ij}}$
$\mathrm{GGo}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}, \nu\right)$	$\mathrm{Ga}\left(0, \frac{1}{\beta}, \alpha\right)$	$\beta_{ij} = \beta_{i,j-1} + \mathrm{e}^{\kappa z_{ij}^{\nu}}$
$\mathrm{S}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{Ga}\left(0, \frac{1}{\beta}, \alpha\right)$	$\beta_{ij} = \beta_{i,j-1} + G(z_{ij})^{\kappa}$
$\mathrm{E}\left(0, \frac{1}{\lambda}\right)$	$\mathrm{IG}\left(1, \frac{1}{\phi}\right)$	$\phi_{ij} = \sqrt{\phi_{i,j-1}\left(\phi_{i,j-1} + 2z_{ij}\right)}$
$\mathrm{W}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{IG}\left(1, \frac{1}{\phi}\right)$	$\phi_{ij} = \sqrt{\phi_{i,j-1}\left(\phi_{i,j-1} + 2z_{ij}^{\kappa}\right)}$
$\mathrm{Go}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}\right)$	$\mathrm{IG}\left(1, \frac{1}{\phi}\right)$	$\phi_{ij} = \sqrt{\phi_{i,j-1}\left(\phi_{i,j-1} + 2\mathrm{e}^{\kappa z_{ij}}\right)}$
$\mathrm{GGo}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}, \nu\right)$	$\mathrm{IG}\left(1, \frac{1}{\phi}\right)$	$\phi_{ij} = \sqrt{\phi_{i,j-1}\left(\phi_{i,j-1} + 2\mathrm{e}^{\kappa z_{ij}^{\nu}}\right)}$
$\mathrm{S}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{IG}\left(1, \frac{1}{\phi}\right)$	$\phi_{ij} = \sqrt{\phi_{i,j-1}\left(\phi_{i,j-1} + 2G(z_{ij})^{\kappa}\right)}$
$\mathrm{E}\left(0, \frac{1}{\lambda}\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2z_{ij}$
$\mathrm{W}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2z_{ij}^{\kappa}$
$\mathrm{Go}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2\mathrm{e}^{\kappa z_{ij}}$
$\mathrm{GGo}\left(\frac{-\ln[\lambda]}{\kappa}, \frac{1}{\kappa}, \nu\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2\mathrm{e}^{\kappa z_{ij}^{\nu}}$
$\mathrm{S}\left(0, \frac{1}{\lambda}, \kappa\right)$	$\mathrm{GIG}\left(\sqrt{\frac{\psi}{\phi}}, \frac{1}{\phi}, \epsilon\right)$	$\phi_{ij} = \phi_{i,j-1} + 2G(z_{ij})^{\kappa}$

Here $\mathrm{W}(\text{location}, \text{scale}, \text{shape})$ is a Weibull distribution, $\mathrm{Go}(\text{location}, \text{scale})$ is a Gompertz distribution, $\mathrm{GGo}(\text{location}, \text{scale}, \text{family})$ is a generalized Gompertz distribution, and $G(y)$ is a function of y determining the distribution $\mathrm{S}(\text{location}, \text{parameter 1}, \text{parameter 2})$.

The corresponding conditional mixture cdf is then

$$F_{\mathrm{m}}(Z_{ij} = k_j | H_{i,j-1}) = 1 - \left(\frac{\beta_{i,j-1}}{\beta_{i,j-1} - \log\left[1 - p_{ijk_j}\right]}\right)^{\alpha} \tag{5.23}$$

The conditional probabilities are obtained by taking the appropriate differences of the conditional cumulative probability, corresponding to the cut-off point of each response category. Hence, the conditional probability of observation j from individual i being in category k is

$$
\begin{aligned}
\Pr(Z_{ij}=0|H_{i,j-1}) &= 1-\left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log[1-p_{ij0}]}\right)^{\alpha} \\
\Pr(Z_{ij}=r_j|H_{i,j-1}) &= \left[1-\left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log\left[1-p_{ijr_j}\right]}\right)^{\alpha}\right] \\
&\qquad -\left[1-\left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log\left[1-p_{ij,r_j-1}\right]}\right)^{\alpha}\right] \\
&= \left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log\left[1-p_{ij,r_j-1}\right]}\right)^{\alpha} \\
&\qquad -\left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log\left[1-p_{ijr_j}\right]}\right)^{\alpha} \\
\Pr(Z_{ij}=\mathtt{nc}-1|H_{i,j-1}) &= \left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}-\log[1-p_{ij,\mathtt{nc}-1}]}\right)^{\alpha}
\end{aligned}
$$

where $0<r_j<(\mathtt{nc}-1)$.

The multivariate density obtained by applying Equation (5.21) consists of products of differences. Unfortunately, these do not simplify as for continuous responses. It is therefore not possible to relate dynamic models obtained using this technique directly to Archimedean (survival) copulas.

Updating equations In this case, the set of updates provided by Bayes's formula also appears to be the most logical choice. But many other updates are possible because the conditional probabilities of Equation (5.21) are independent.

PROPOSITION 5.6

An interesting case occurs when the corresponding density can also be obtained as a (closed under sampling) mixture and the set of frailty updates deducted is used. It can directly be noticed that this only provides a second updating equation. But in fact, the conditional mixture cdfs (before taking differences) now correspond to the conditional mixture densities which would be obtained in the case of continuous responses. Indeed,

$$
\frac{\partial F_{\mathrm{m}}(z_{ij}|H_{i,j-1})}{\partial z_{ij}} = f_{\mathrm{m}}(z_{ij}|H_{i,j-1}) \tag{5.24}
$$

as subject i's history $H_{i,j-1}$ is a constant with respect to z_{ij} and the updating equations do not change the form of the conditional mixtures (due to the closed under sampling or recursively closed property). Hence, using the set of frailty updates asymptotically tends to a frailty model corresponding to an Archimedean (survival) copula as the number of categories increases towards infinity. ○

As an illustration, consider the case of conditional gamma mixture cdf presented in Equation (5.23) with the set of frailty updates given in Equation (5.9). The conditional mixture cdf is

$$F_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = 1 - \left(\frac{\beta_{i,j-1}}{\beta_{i,j-1}+z_{ij}}\right)^{\alpha} = 1 - \left(\frac{\frac{1}{\delta}+\sum_{r=1}^{j-1} z_{ir}}{\frac{1}{\delta}+\sum_{r=1}^{j} z_{ir}}\right)^{\frac{1}{\delta}+j-1}$$

and the conditional mixture density is

$$f_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = \frac{\alpha_{i,j-1}\beta_{i,j-1}^{\alpha_{i,j-1}}}{(\beta_{i,j-1}+z_{ij})^{\alpha_{i,j-1}+1}} = \frac{\left(\frac{1}{\delta}+j-1\right)\left(\frac{1}{\delta}+\sum_{r=1}^{j-1} z_{ir}\right)^{\frac{1}{\delta}+j-1}}{\left(\frac{1}{\delta}+\sum_{r=1}^{j} z_{ir}\right)^{\frac{1}{\delta}+j}}$$

where $\alpha_{i0} = \beta_{i0} = \frac{1}{\delta}$ (recall that α_{i0} and β_{i0} are not identifiable in the initial state). Thus, the relationship of Equation (5.24) does hold.

Obviously, the serial and Markov updates respectively given in Equations (5.11) and (5.12) also provide possible alternative sets of updating equations.

5.3.2 PRODUCT OF CONDITIONAL CDFS APPROACH

In this approach, no hypotheses are made about the discrete process. Hence, the appropriate differences in m dimensions of the multivariate cdf (see Section 4.1) are used to construct the likelihood. As previously, the arbitrary function $G(y)$ inserted into the mixture will involve the cdf (or survival function) of one of the (nominal or ordinal) parametrizations of Section 1.3 and be denoted by z.

Recall that all conditional mixture cdfs are chosen to be the same. Hence, all the mixture cdfs (see Section 3.1) and mixture survival functions (see Section 3.2) must be recursively closed. They are respectively shown in Tables 5.2 and 5.3.

From Equation (5.3), it can be seen that an appropriate multivariate cdf is not usually produced from the product

$$F_{\mathrm{m}}(z_{i1})F_{\mathrm{m}}(z_{i2}|H_1)...F_{\mathrm{m}}(z_{im}|H_{m-1})$$

where $\frac{\partial F_{\mathrm{m}}(z_{ij}|H_{j-1})}{\partial z_{ij}} = f_{\mathrm{m}}(z_{ij}|H_{j-1})$. Hence, the multivariate cdf is constructed from the product

$$\begin{aligned} F_{\mathrm{M}}(z_{i1},\dots,z_{im}) &= F(z_{i1})\frac{F_{\mathrm{M}}(z_{i1},z_{21})}{F_{\mathrm{M}}(z_{i1})}\cdots\frac{F_{\mathrm{M}}(z_{i1},\dots,z_{im})}{F_{\mathrm{M}}(z_{i1},\dots,z_{i,m-1})} \quad (5.25)\\ &= F(z_{i1})F^{*}_{\mathrm{m}}(z_{i2}|H_1)\dots F^{*}_{\mathrm{m}}(z_{i,m}|H_{m-1}) \end{aligned}$$

where the first term of the series is $F(z_{i1})$ and the conditional cdfs $F^{*}_{\mathrm{m}}(z_{ij}|H_{j-1})$ correspond to the marginal ratios $\frac{F_{\mathrm{M}}(z_{i1},\dots,z_{ij})}{F_{\mathrm{M}}(z_{i1},\dots,z_{i,j-1})}$.

Unfortunately, mixtures recursively closed do not always correspond to this ratio once updated.

PROPOSITION 5.7

A recursively closed mixture corresponds to the ratio $\frac{F_{\mathrm{M}}(z_{i1},\dots,z_{ij})}{F_{\mathrm{M}}(z_{i1},\dots,z_{i,j-1})}$ once updated, if the following conditions are satisfied.

- The mixture is a ratio where the numerator and/or the denominator involves some parameter, say δ, corresponding to the update equation which is only added to (or subtracted from) the response or z_{ij}.
- This ratio is equal to a constant (with respect to the parameter δ) when the response or z_{ij} is set equal to zero (the numerator is then equal to the denominator).

Hence, the updating equation is $\delta_{ij} = \delta_{i,j-1} + z_{ij}$ and the ratio has general form

$$c_{j1}\left[\frac{g(\delta_{ij})}{g(\delta_{i,j-1})}\right]^{c_2} \tag{5.26}$$

where c_{j1} and c_2 are constants with respect to δ and $g(\cdot)$ is an arbitrary function. *Proof:* The denominator is one step behind the numerator. Each numerator must cancel with the denominator at the previous step, see Equation (5.25). Any constants (c_2) acting as a power or any arbitrary function ($g(\cdot)$) will automatically cancel out as long as they are applied to both the numerator and denominator. All the constants acting multiplicatively (c_{j1}) remain as normalizing constants of the resulting multivariate distribution in addition to the normalizing constant of the first unconditional term. Hence, the ratio can be simplified to

$$\frac{\delta_{ij}}{\delta_{i,j-1}} = \frac{\delta_{i,j-1} + z_{ij}}{\delta_{i,j-1}}$$

where the numerator will indeed cancel with the denominator at the previous step. From this, it can also be seen that the resulting multivariate distribution will corresponds to a multivariate Laplace transform (Subsection 2.2.3) or an Archimedean (survival) copula with mixture margins (Subsection 4.3.2) because it has general form

$$\left[\frac{g\left(\delta_{i0} + \sum_{j=1}^{m} z_{ij}\right)}{g(\delta_{i0})}\right]^{c_2} \prod_{j=1}^{m} c_{j1}$$

where the updating parameter is defined as δ_{i0} in the initial state. ○

Examples In the case of an exponential survival function, Equation (B.6), with a one parameter exponential density, Equation (B.5), the distribution integrated out is also a one parameter exponential distribution (Subsection 3.2.1)

$$\begin{aligned} f(\lambda_{ij}|H_{ij}) &= [\beta_{i,j-1} + z_{ij}]\mathrm{e}^{-\lambda_{ij}[\beta_{i,j-1}+z_{ij}]} \\ &= \beta_{ij}\mathrm{e}^{-\beta_{ij}\lambda_{ij}} \end{aligned} \tag{5.27}$$

and the mixture survival function is

$$S_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = \frac{\beta_{i,j-1}}{\beta_{i,j-1} + z_{ij}} \tag{5.28}$$

The updating equation can now be deduced from the distribution integrated out, Equations (5.27).

$$\beta_{ij} = \beta_{i,j-1} + z_{ij}$$

The mixture survival function presented in Equation (5.28) matches the general form presented in Equation (5.26) where c_{j1} is unity, $g(\cdot)$ is the identity function, and c_2 is minus one. Thus, this will result in a proper multivariate survival function of the form

$$S_{\mathrm{M}}(\mathbf{z}) = \left[1 + \frac{\sum_{j=1}^{m} z_{ij}}{\delta_{i0}}\right]^{-1}$$

The complementary log-log link given in Equation (1.9) corresponds to the cdf of a Gompertz distribution. Hence, the proportional hazard model of McCullagh (1980) presented in Equation (1.10) is the integrated hazard of a Gompertz distribution. This can therefore be chosen as the z_{ij}. Then the multivariate survival function is

$$S_{\mathrm{M}}(Z_{i1} = k_1, \ldots, Z_{im} = k_m) = \left[1 + \frac{\sum_{j=1}^{m} \theta_{k_j} \mathrm{e}^{-\sum_{l=1}^{\mathtt{ncv}} \phi_l x_{ijl}}}{\delta_{i0}}\right]^{-1} \qquad (5.29)$$

where θ_{k_j} and ϕ_l are respectively the intercept and (time-varying) covariate (x_{ijl}) regression parameters. Hence, z_{ij} corresponds to the integrated hazard of a Gompertz distribution, Equation (B.21).

Because each univariate marginal survival function has the form of a univariate generalized logistic distribution, Equation (B.33),

$$S(Z_{ij} = k_j) = \left[1 + \frac{\mathrm{e}^{\alpha_{k_j} + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{ijl}}}{\delta_{i0}}\right]^{-1}$$

this multivariate survival function can be identified as a multivariate generalized logistic distribution.

Note that this multivariate distribution does not correspond to the multivariate logistic distribution presented in Equation (B.51) when δ_{i0} is fixed to unity. Indeed, care must be taken because constructing a multivariate cdf and a multivariate survival function will produce two different multivariate distributions.

For instance, the log-log link given in Equation (1.11) corresponds to the survival function of a Gompertz distribution. The equivalent of the proportional hazards model presented in Equation (1.10) can then be obtained

$$\begin{aligned} -\ln(\mu_i) &= \mathrm{e}^{\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}} \qquad (5.30) \\ &= \theta_k \mathrm{e}^{-\sum_{l=1}^{\mathtt{ncv}} \phi_l x_{il}} \end{aligned}$$

where $\theta_k = \mathrm{e}^{\alpha_k}$ and $\phi = -\beta$.

By choosing Equation (5.30) for the z_{ij}, the range of the multivariate distribution is inverted resulting in a multivariate cdf which corresponds to the associated copula.

This first illustration was quite simple because the dependence among observations was fixed. Indeed, this can be seen to be a special case of a gamma Archimedean survival copula where the dependence is fixed at unity (note that this does not correspond to independence).

In this second example, an exponential survival function, Equation (B.6), is combined with a generalized inverse Gaussian, Equation(B.10), mixing distribution,

$$p(\lambda_{ij}|H_{i,j-1}) = \frac{\lambda_{ij}^{\epsilon-1} \mathrm{e}^{-\frac{\lambda_{ij}\phi_{i,j-1}}{2} - \frac{\psi}{2\lambda_{ij}}}}{2\left(\frac{\psi}{\phi_{i,j-1}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi_{i,j-1}}, |\epsilon|\right)}$$

the distribution integrated out is also a one parameter exponential distribution (Subsection 3.2.4)

$$\begin{aligned} f(\lambda_{ij}|H_{ij}) &= \frac{\lambda_{ij}^{\epsilon-1} \mathrm{e}^{-\frac{\lambda_{ij}}{2}(\phi_{i,j-1}+2z_{ij}) - \frac{\psi}{2\lambda_{ij}}}}{2\left(\frac{\psi}{\phi_{i,j-1}+2z_{ij}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi(\phi_{i,j-1}+2z_{ij})}, |\epsilon|\right)} \\ &= \frac{\lambda_{ij}^{\epsilon-1} \mathrm{e}^{-\frac{\phi_{ij}\lambda_{ij}}{2} - \frac{\psi}{2\lambda_{ij}}}}{2\left(\frac{\psi}{\phi_{ij}}\right)^{\frac{\epsilon}{2}} K\left(\sqrt{\psi\phi_{ij}}, |\epsilon|\right)} \end{aligned} \tag{5.31}$$

and the mixture survival function obtained is

$$S_{\mathrm{m}}(z_{ij}|H_{i,j-1}) = \frac{K\left(\sqrt{\psi(\phi_{i,j-1}+2z_{ij})}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi_{i,j-1}}, |\epsilon|\right)} \left(\frac{\phi_{i,j-1}}{\phi_{i,j-1}+2z_{ij}}\right)^{\frac{\epsilon}{2}} \tag{5.32}$$

The updating equation can now be deduced from the distribution integrated out, Equations (5.31).

$$\phi_{ij} = \phi_{i,j-1} + 2z_{ij}$$

The mixture survival function presented in Equation (5.32) matches the general form presented in Equation (5.26) where c_{j1} is unity, c_2 is $-\frac{\epsilon}{2}$, and the function $g(\cdot)$ is separated in two parts

$$g_1(x) = K\left(\sqrt{x\psi}, |\epsilon|\right)$$

and $g_2(\cdot)$, the identity function. Thus, this will result in a proper multivariate survival function of the form

$$S_{\mathrm{M}}(\mathbf{z}) = \frac{K\left(\sqrt{\psi\left(\delta_{i0}+2\sum_{j=1}^{m} z_{ij}\right)}, |\epsilon|\right)}{K\left(\sqrt{\psi\delta_{i0}}, |\epsilon|\right)} \left(1+\frac{2}{\delta_{i0}}\sum_{j=1}^{m} z_{ij}\right)^{-\frac{\epsilon}{2}} \tag{5.33}$$

Hougaard (2000, p. 246) considers that this model can easily be obtained as the generalized inverse Gaussian belongs to the exponential family.

If the responses are ordinal and the proportional odds parametrization of Subsection 1.3.5 is used with a Weibull link function then the multivariate survival function obtained is

$$S_{\mathrm{M}}(Z_{i1}=k_1,\ldots,Z_{im}=k_m) = \frac{K\left(\sqrt{\psi\left[\delta_{i0}+2\sum_{j=1}^{m}(\alpha_{k_j}+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{ijl})^{\kappa}\right]}, |\epsilon|\right)}{K\left(\sqrt{\psi\delta_{i0}}, |\epsilon|\right)} \left[1+\frac{2}{\delta_{i0}}\sum_{j=1}^{m}\left(\alpha_{k_j}+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{ijl}\right)^{\kappa}\right]^{-\frac{\epsilon}{2}} \tag{5.34}$$

where α_{k_j} and β_l are respectively the intercept and (time-varying) covariate (x_{ijl}) regression parameters and z_{ij} corresponds to the integrated hazard of a Weibull distribution, Equation (B.14).

Note that each univariate marginal survival function is a mixture distribution of the form

$$S_{\mathrm{m}}(Z_{ij}=k_j) = \frac{K\left(\sqrt{\psi\left[\phi_{i,j-1}+2\left(\alpha_{k_j}+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{ijl}\right)^{\kappa}\right]}, |\epsilon|\right)}{K\left(\sqrt{\psi\phi_{i,j-1}}, |\epsilon|\right)} \left[\frac{\phi_{i,j-1}}{\phi_{i,j-1}+2\left(\alpha_{k_j}+\sum_{l=1}^{\mathtt{ncv}}\beta_l x_{ijl}\right)^{\kappa}}\right]^{\frac{\epsilon}{2}}$$

and has yet to be named as the multivariate distribution obtained. Hougaard (2000, p. 352) confirms that models obtained from Equation (5.33) have not been examined in detail.

The likelihood corresponding to Equations (5.29) and (5.34) can now be obtained by taking the appropriate differences as described in Section 4.1.

Finally, it is important to remember that the multivariate distribution corresponding to the associated copula can always be considered as an alternative model. Recall that although the respective multivariate cdf and multivariate survival functions have a similar form, they do not correspond to the same multivariate distribution.

Updating equations Although all the mixtures in Tables 5.2 and 5.3 satisfy the conditions stated in Proposition 5.7, the latter can be adapted for multiplicative updates.

PROPOSITION 5.8
A recursively closed mixture corresponds to the ratio $\frac{F_{\mathrm{M}}(z_{i1},\ldots,z_{ij})}{F_{\mathrm{M}}(z_{i1},\ldots,z_{i,j-1})}$ once updated, if the following conditions are satisfied.

- The mixture is a ratio where the numerator and/or the denominator involves some parameter, say δ, corresponding to the update equation which is only multiplied (or divided) by the response or z_{ij}.
- This ratio is equal to a constant (with respect to the parameter δ) when the response or z_{ij} is set equal to unity (the numerator is then equal to the denominator).

Hence, the updating equation is $\delta_{ij} = \delta_{i,j-1} z_{ij}$ and the ratio has general form

$$c_{j1} \left[\frac{g(\delta_{ij})}{g(\delta_{i,j-1})} \right]^{c_2}$$

where c_{j1} and c_2 are constants with respect to δ and $g(\cdot)$ is an arbitrary function. The resulting multivariate distribution will corresponds to an Archimedean (survival) copula, constructed from a multiplicative generator, with mixture margins (Subsection 4.3.2) because it has the general form

$$\left[\frac{g\left(\delta_{i0} \prod_{j=1}^{m} z_{ij}\right)}{g(\delta_{i0})} \right]^{c_2} \prod_{j=1}^{m} c_{j1}$$

where the updating parameter is defined as δ_{i0} in the initial state. ○

Propositions 5.7 and 5.8 can both be adapted for continuous responses by splitting the constants into two parts to allow for the changes brought by the second updating equation (and writing it in terms of densities). It can now be better understood why the Gaussian mixture obtained in Subsection 3.3.1 cannot be updated to correspond to a frailty model linked to an Archimedean (survival) copula.

Linking conditional to marginal models The conditional cdf or survival distribution may not necessarily be a mixture distribution. For instance, the power variance Laplace transform (Hougaard, 1986a; Hougaard, 2000, pp. 504–506)

$$\phi(s) = \mathrm{e}^{-\frac{\delta}{\alpha}[(\theta+s)^{\alpha} - \theta^{\alpha}]}$$

obtained from the moment generating function also satisfies the conditions of Proposition 5.7. It therefore corresponds to the general form presented in Equation (5.26) where the function $g(\cdot)$ is

$$g(x) = \mathrm{e}^{-\frac{\delta}{\alpha} x^{\alpha}}$$

and c_{j1} and c_2 are fixed at unity.

The power variance Archimedean copula (Hougaard, 2000, pp. 241–243) can be written as

$$\mathrm{C_M}(\mathbf{u}) = \mathrm{e}^{-\frac{\delta}{\alpha}\left(\left[\sum_{j=1}^{m}\left(\theta^{\alpha}-\frac{\alpha}{\delta}\ln[u_j]\right)^{\frac{1}{\alpha}}-(m-1)\theta\right]^{\alpha}-\theta^{\alpha}\right)}$$

where u_j is a $U(0,1)$ cdf for margin j, $\mathbf{u}$ is a vector containing the univariate margins and m represents the number of univariate margins (see Appendix C.3).

Crouchley (1995) has used this copula to model ordinal data. The univariate marginal distributions were chosen to be power variance functions with the z_{ij} chosen as the integrated intensity of a Gompertz distribution, Equation (B.21). Recall from Equation (5.30) that this is equivalent to a proportional odds model (Subsection 1.3.5) with a log-log link.

The corresponding multivariate cdf was used to model the ordinal responses and is given for the bivariate and the trivariate cases. In the bivariate case, this is similar to the relationship linking copulas and associated copulas (see Property 2.18). This can be generalized to m dimension (Joe, 1997, p. 10) as

$$F_{\mathrm{M}}(\mathbf{x}) = 1 + \sum_{r=1}^{m}(-1)^r \sum_{j=1}^{\binom{m}{r}} S_{\mathrm{R}}(\mathbf{x}_{\Theta_j}) \tag{5.35}$$

where $S_{\mathrm{R}}(\mathbf{x}_{\Theta_j})$ is a univariate survival function if r is unity and otherwise a joint survival function of dimension r. The subscript Θ_j denotes the margins involved. Hence in the trivariate case, the relationship is

$$\begin{aligned} F_{\mathrm{T}}(x_1,x_2,x_3) = {} & 1 - S(x_1) - S(x_2) - S(x_3) \\ & + S_{\mathrm{B}}(x_1,x_2) + S_{\mathrm{B}}(x_1,x_3) + S_{\mathrm{B}}(x_2,x_3) - S_{\mathrm{T}}(x_1,x_2,x_3) \end{aligned}$$

The likelihood is then obtained by taking the appropriate differences as described in Section 4.1.

Crouchley (1995) also considers the special cases where the power variance Laplace transform (see Appendix C.3) reduces to a positive stable and a gamma Laplace transform. The later special case has also been considered by several other people.

Conaway (1990) used the gamma survival copula to model clustered binary data. The multivariate distribution was obtained from a mixture combining a Gompertz survival function, Equation (B.20), and a log-gamma mixing density which was then integrating with respect to the mixing distribution. Using the present method, this is equivalent to inserting the proportional hazard model of McCullagh (1980) presented in Equation (1.10) in a gamma Laplace transform. Hence, the univariate marginal distributions were gamma mixtures and the binary parametrization of Subsection 1.3.1 was used with the log-log link. In addition, serial dependence was added in the model by introducing the lagged response to the regression model. A first order Markov model was also considered by imposing certain constraints among these regression coefficients.

Finally, Ten Have (1996) also uses the special case of the gamma survival copula where gamma mixture margins are inserted (Subsection 4.3.2) with the z_{ij} chosen as the integrated intensity of a Gompertz distribution, Equation (B.21). The model is used to analyze ordinal responses in several dimensions and extended to include discrete failure times with ordinal responses. Pulkstenis *et al.* (1998) apply this specific model to binary data where the dropout process is also taken into account.

5.3.3 RECURSIVE PROBABILITIES

5.4 Hidden Markov chains

The dependence structure in dynamically updated models was constructed conditionally by taking each subject's previous observed history into account. Dependence among each subject's observations will now be induced differently for the case where the responses are generated in several different unknown states. In such a case, it is assumed that subjects change over time among a number of different underlying (hidden) processes, called states. As the subjects are not restricted to be influenced by the same set of external factors (covariates) in each of these states, a (linear or non-linear) regression is required for each one. Obviously, the model can be simplified by allowing common parameters among these different regression equations. Now, the dependence is induced conditional on the subject's previous state history rather than its previous observed history. This type of dependence can be called spells.

All possible changes of state over time must be taken into account by the probability model as they are not observed. Each subject's possible "path" through the states corresponds to a product of conditional probabilities. Because the dependence is conditioning on each subject's previous state history, it is calculated by multiplying the probability of a particular subject at a specific time point by what is called a transition probability. Hence, the transition probabilities are dependencies measuring the probabilities of changing state (or remaining in it). This ensures (as they are probabilities) that summing all these products of conditional probabilities together produces the joint probability over time and all possible states for a particular subject. The sample probability is then obtained by multiplying these sum of products of conditional probabilities together over all individuals, yielding a likelihood.

One last problem remains concerning the initial conditions. At the first time point, no previous information is available to estimate the probability of being in a particular state. This can be solved by either introducing $s-1$ additional parameters for s states or by using the marginal transition probabilities. This latter case assumes that stationarity of the hidden process has been reached. This solution is chosen and has $s(s-1)$ transition probabilities to be estimated along with the regression parameters. Note that the probability of remaining in state r is equal to $1-\sum_{o=1}^{s-1}\gamma_{or}$ where γ_{or} are the transition probabilities.

Consider a simple case with two states and three time points. The transition

	State 1	State 2
State 1	$1-\gamma_1$	γ_1
State 2	γ_2	$1-\gamma_2$

Fig. 5.1. Transition probabilities for a two state model.

probabilities are shown in Figure 5.1 and the marginal probabilities are $\delta = \frac{\gamma_2}{\gamma_1+\gamma_2}$ and $1-\delta$ respectively for state 1 and state 2 (MacDonald and Zucchini, 1997, pp. 67–68). The joint probability over time and all possible states for subject i is then

$$
\begin{aligned}
&\Pr(Y_{i1}=k_1,\ldots,Y_{im}=k_m) = \\
&\quad \delta \Big\{ \Pr(Y_{i1}=k_1|S_{11})(1-\gamma_1)\Big[\Pr(Y_{i2}=k_2|S_{21})\ (1-\gamma_1)\ \Pr(Y_{i3}=k_3|S_{31}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\gamma_1 \ \Pr(Y_{i3}=k_3|S_{32})\Big] \\
&\qquad\qquad\qquad\qquad\quad +\gamma_1 \Big[\Pr(Y_{i2}=k_2|S_{22})\ \ \gamma_2\ \ \Pr(Y_{i3}=k_3|S_{31}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +(1-\gamma_2)\Pr(Y_{i3}=k_3|S_{32})\Big]\Big\} \\
&+(1-\delta)\Big\{ \Pr(Y_{i1}=k_1|S_{12})\ \ +\gamma_2 \Big[\Pr(Y_{i2}=k_2|S_{21})\ (1-\gamma_1)\ \Pr(Y_{i3}=k_3|S_{31}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\gamma_1 \ \Pr(Y_{i3}=k_3|S_{32})\Big] \\
&\qquad\qquad\qquad\qquad\quad (1-\gamma_2)\Big[\Pr(Y_{i2}=k_2|S_{22})\ \ \gamma_2\ \ \Pr(Y_{i3}=k_3|S_{31}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +(1-\gamma_2)\Pr(Y_{i3}=k_3|S_{32})\Big]\Big\}
\end{aligned}
$$

where $\Pr(Y_{ij}=k_j|S_{jh})$ is the probability of the response being observed in category k at time j for subject i given it is in state h at time j.

Notice that this collapses to the product of independent probabilities if there is only one state.

Unfortunately, the number of terms expands exponentially as the number of time points increases and quadratically for the number of states considered. This expression is therefore not computationally feasible as it stands for any reasonable number of states and/or time points. However, it can be rearranged in a recursive form over time (MacDonald and Zucchini, 1997, pp. 78–79; Lindsey, 1999, p. 73). The joint probability over time and all possible states for subject i can then be written as

$$\Pr(Y_{i1} = k_1, \ldots, Y_{im} = k_m) = \boldsymbol{\delta}^T \prod_{j=1}^{m} (\boldsymbol{\gamma}\boldsymbol{\pi}_{ijk_j})\mathbf{J}^T \tag{5.36}$$

where $\boldsymbol{\delta}$ is a row vector containing the marginal probabilities, $\boldsymbol{\gamma}$ is the transition matrix, $\boldsymbol{\pi}_{ijk_j}$ is an $s \times s$ matrix containing on the diagonal the probabilities of the response being observed in category k at time j for subject i given the various possible states, and $\mathbf{J}$ is a row vector of ones.

The likelihood can then be obtained by multiplying Equation (5.36) over all subjects. Note that at the first time point the Equation (5.36) reduces to

$$\Pr(Y_{i1} =) = k_1 \boldsymbol{\delta}^T \boldsymbol{\pi}_{ijk_j} \mathbf{J}^T$$

as the vector of marginal probabilities $\boldsymbol{\delta}$ is obtained by solving $\boldsymbol{\delta}^T(\mathbf{I} - \boldsymbol{\gamma}) = 0$.

In the two state and three time point example above, the vector of marginal probabilities is

$$\boldsymbol{\delta} = \begin{pmatrix} \frac{\gamma_2}{\gamma_1 + \gamma_2} & \frac{\gamma_1}{\gamma_1 + \gamma_2} \end{pmatrix}$$

and the transition matrix is

$$\boldsymbol{\gamma} = \begin{pmatrix} 1 - \gamma_1 & \gamma_1 \\ \gamma_2 & 1 - \gamma_2 \end{pmatrix}$$

as presented in Figure 5.1.

Finally, as the responses are nominal or ordinal, the probabilities at time j contained on the diagonal of $\boldsymbol{\pi}_j$ are obtained using one of the parametrizations presented in Section 1.3.

Filtered and recursive probabilities Two types of recursive probabilities can now be extracted from this model.

These are obtained from the intermediate values

$$\zeta_{ijr} = \sum_{o=1}^{s} \zeta_{i,j-1,o} \gamma_{or} \Pr(Y_{ij} = k_j | S_{jr})$$

calculated while constructing the joint probability over time and all possible states for subject i are required, where

$$\zeta_{i1r} = \delta_r \Pr(Y_{i1} = k_1 | S_{1r})$$

is obtained for the first time point.

A filtered probability is the probability that a specific subject is in a particular hidden state given this subject's previous state history. Hence, the probability of subject i being in state r at time j is

$$\xi_{ijr} = \frac{\zeta_{ijr}}{\sum_{o=1}^{s} \zeta_{ijr}}$$

which is obtained by standardizing the ζ_{ijr}.

The probabilities of the response being observed in category k at time j for subject i can then be calculated by

$$\varphi_{ij} = \sum_{o=1}^{s} \xi_{ijo} \,\mathrm{Pr}(Y_{ij} = k_j | S_{jo})$$

which is the recursive probabilities for subject i.

5.5 Further reading

6
Graphical representation of longitudinal ordinal data

6.1 Visualizing the collected data

In contrast to continuous longitudinal responses, an individual profile plot of longitudinal ordinal data is not very informative. Indeed, due to the few different possible levels of outcome, most profiles overlap. Hence, a different type of graph such as a plot of the cumulative probabilities over time will be more informative.

Such plots are quite different from individual profile plots as they do not represent the data at an individual level. Indeed, the outcomes of a particular individual cannot be followed over time in this way. On the other hand, they give a general idea of changes in time of the patients' distribution over the possible outcomes.

The cumulative probabilities over time can also be plotted separately for various groups.

Cumulative probability plots are useful to picture ordinal responses as these are directly modelled by the regression curve.

6.2 Visualizing predicted probabilities

Because a dynamic dependence has been created among the repeated observations, two totally different types of curves can be obtained: individual or recursive fitted values and population or average predictions. Indeed, a recursive curve is fundamentally different from an average curve as it is not representative of the entire population but of a particular subject. In other words, a specific recursive curve represents all individuals who not only have identical covariate patterns but who would also respond to the covariates in exactly the same way over time. This is due to the dynamic part of the model which readjusts the recursive curve at each time point accordingly to the previous response of the person under consideration. Hence, recursive curves are necessarily individual specific.

Nevertheless, both types of curves can be represented by cumulative probabilities or highest probability categories. A recursive mean curve can also be computed but is less informative if the underlying scale has not been respected. In other words, if the categories do not correspond to identical ranges of the underlying scale then the recursive mean curves will be distorted.

Thus, a dynamic model for ordinal responses with a significant dependence can be represented by at least five distinct curves: two population and three recursive curves.

6.2.1 UNDERLYING MEAN PROFILES

6.2.2 RECURSIVE OR INDIVIDUAL PROFILES

The computed recursive means yield smooth curves typical of a cumulative dependence. The dynamic process can clearly be observed from these curves. Indeed, all recursive mean curves follow the underlying population curve but are slightly pulled towards a patient's observations. The recursive highest probabilities can be very useful as they represent a specific response category. These curves follow each patient's observations reasonable well.

6.2.3 TIME SPENT IN A HIDDEN STATE

7
Data analysis and model comparison

As the modelling process is exploratory, the inference criterion used for comparing the models under consideration is their ability to predict the observed data, that is how probable they make the data. In other words, models are compared directly through their minimized minus log likelihood. When the numbers of parameters in models differ, they are penalized by adding the number of estimated parameters, a form of the Akaike information criterion (AIC, see Akaike, 1973; Lindsey and Jones, 1998). Smaller values indicate more preferable models. This criterion allows direct comparisons among models, that are not required to be nested.

7.1 The heart examination clinical trial

This clinical trial involved patients requiring a heart examination. Doctors are able to assess the left ventricular volume and ejection fraction for each patient by intermittent harmonic colour Doppler. This is only possible if each patient remains within a particular range of a Doppler signal induced by a drug administered by intravenous infusion. The aim of the trial is to study the relationship between the concentration of the drug administered and the amount of time the patients spent within the required signal range.

The trial involves 85 patients randomized into three concentration groups: 17 receiving 2.5g, 30 receiving 4g, and 38 receiving 8g of the drug under study. The Doppler signal of each patient was recorded continuously over time. This was then discretized into four ordered categories, namely no signal, an insufficient signal, an appropriate signal, and an excess signal. These were respectively coded as levels 0 (baseline), 1, 2, and 3.

The signal was discretized at frequent regular intervals over time resulting in a total of 19690 observations being collected. Indeed, all patients were categorized at the same time points but have series of different lengths. This is due to different rates of drug intake and wearing off among the patients. The series have a maximum length of 781 seconds for the 2.5g group, 2089 for the 4g group, and 1643 for the 8g group.

The responses are also unbalanced with respect to the chosen discretization time points. These were chosen very close together in time (a few seconds apart) at the start of the trial, slightly further apart (up to thirty seconds) during the period where an appropriate signal level was observed, and finally far apart (up to sixty seconds or more) when the drug effect started wearing off.

The intravenous infusion rate started at 1ml per minute. Depending on the patient's response to the drug, this infusion rate could either be decreased to 0.5ml per minute, or increased to 2 or 4ml per minute. This was necessary in a few cases to obtain or maintain a particular patient at the appropriate signal level.

The data are shown in this way in Figure 7.1. The change over time of the probability of being in a particular category is represented in Figure 7.1a. Most patients have no Doppler signal during the first 50 seconds of the trial. Then, patients have a rapid increase in signal level. During the next 50 seconds patients have an insufficient signal level. After 100 seconds, the majority of the patients remain at an appropriate signal level for almost 13 minutes. From then on, a patient's signal level slowly disappears.

The cumulative probabilities over time are also plotted separately for the three concentration groups. Patients who had a higher concentration of the drug administered (Figures 7.1c and 7.1d) have longer observed series and also spend a longer time at the appropriate signal level compared to the reference group of 2.5g (Figure 7.1b).

Finally, more variability appears in the two lower concentration groups (2.5g and 4g) once the drug effect starts wearing off (roughly between 500 and 1200 seconds after study initiation).

To begin, an independent multinomial regression was fitted. This null model just contains the three intercept parameters and has an AIC of 19720. This provides us with a reference point for comparison with further fitted models.

Cumulative probability plots are useful to picture ordinal responses as these are directly modelled by the regression curve. From the plots in Figure 7.1, it is clear that the cumulative probabilities have a somewhat parabolic shape over time. For a start, a reasonable regression curve might therefore be a second-degree polynomial in time. This five-parameter proportional odds model, still with independence among observations, lowers the AIC to 16591.

However, it is also clear from Figure 7.1 that the curve is not symmetric. Hence, non-linear regression models must now be considered. Among several models fitted, a sum of exponentials

$$\beta_1 \times \left[\exp\left(1 - \frac{\text{times}}{\beta_2}\right) + \exp\left(\frac{\text{times}}{\beta_3}\right)\right]$$

fits better than the second-degree polynomial. Indeed, this six parameter model lowers the AIC to 14811. This regression curve is similar to a first-order one-compartment model used in pharmacokinetics where β_1 would correspond to a volume parameter, β_2 to the absorption rate, and β_3 to the elimination rate. (All non-linear regression curves are presented without the intercepts; the appropriate intercept for a given ordinal category must be added.)

A model where the flow rate is linearly added to the equation has seven parameters and an AIC of 14669. This model is further improved by also linearly adding

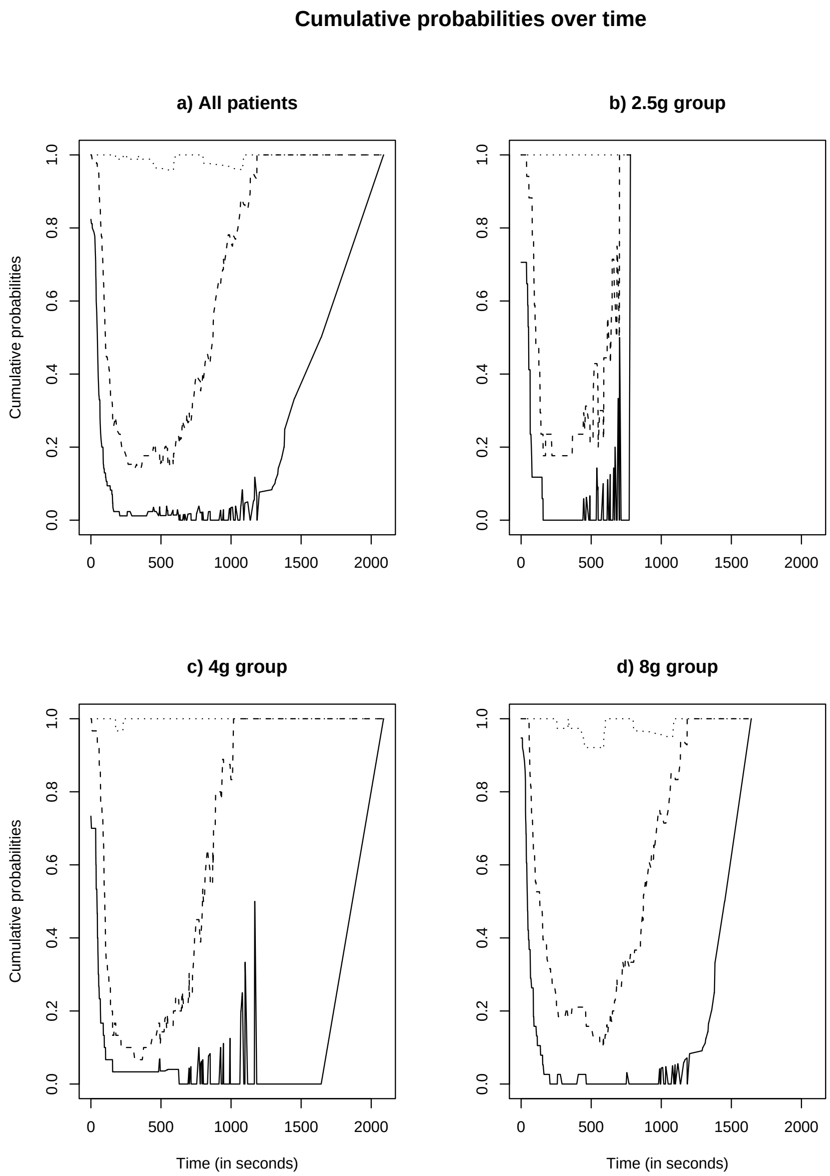

Fig. 7.1. The area below the solid line indicates no signal, that between the solid and the dashed line, an insufficient signal, between the dashed and dotted line, an appropriate signal, and above the dotted line, an excess signal.

the concentration groups, lowering the AIC to 14620 with nine parameters.

Next, the various concentration groups can be allowed to evolve differently over time. This is introduced into the model by creating interactions between the groups and time. Each group requires a different curve. A fourteen-parameter model is fitted which further lowers the AIC to 14448:

$$\beta_1 \times \left[\exp\left(1-\frac{\text{times}}{\beta_2}\right)+\exp\left(\frac{\text{times}}{\beta_3}\right)\right]+\beta_4\times\text{flow}+\beta_5\times\text{group4}+$$
$$\beta_6\times\text{group8}+\beta_7\times\text{group4}\times\left[\exp\left(1-\frac{\text{times}}{\beta_8}\right)+\exp\left(\frac{\text{times}}{\beta_9}\right)\right]+$$
$$\beta_{10}\times\text{group8}\times\exp\left(1-\frac{\text{times}}{\beta_{11}}\right)$$

where β_1 to β_3 describes the change over time for the reference (2.5g) group, β_4 is the infusion rate parameter, β_5 is the 4g group, β_6 is the 8g group, parameters β_7 to β_9 describe the change over time for the 4g group, and β_{10} and β_{11} describe the change over time for the 8g group.

All these models assumed independence among the response observations. An over-dispersion model introduces an additional parameter, but on the other hand, the interaction involving the 4g group can now be simplified. This results in a thirteen-parameter model with a slightly lower AIC of 14441:

$$\beta_1 \times \left[\exp\left(1-\frac{\text{times}}{\beta_2}\right)+\exp\left(\frac{\text{times}}{\beta_3}\right)\right]+\beta_4\times\text{flow}+\beta_5\times\text{group4}+$$
$$\beta_6\times\text{group8}+\beta_7\times\text{group4}\times\exp\left(1-\frac{\text{times}}{\beta_2}\right)+$$
$$\beta_8\times\text{group8}\times\exp\left(1-\frac{\text{times}}{\beta_9}\right)$$

Now, we introduce different types of dependencies among successive observations by applying the gamma Laplace mixture with the proportional hazard model of McCullagh (1980) presented in Equation (1.10) inserted shown as illustration in the latent approach (Subsection 5.3.1). In each case, the regression is identical to that of the over-dispersion model. The cumulative model yields an AIC of 12858 with thirteen parameters, the serial dependence yields an AIC of 10618 with fourteen parameters, and the Markov dependence yields an AIC of 13426 with fourteen parameters, as can be seen in Table 7.1.

The lowest AIC is obtained using the serial update to introduce dependence among the succeeding observations. The dependence parameter is estimated to be 0.95, indicating a very strong dependence on the previous responses. Unfortunately, this dependence close to 1 combined with a particular feature of this trial creates an undesirable effect.

Looking closely at the individual series, we notice that, due to the high frequency of observation, the series consist of long sequences of values at a particular signal level with very few changes. Due to this lack of variability, the high

Table 7.1. AICs for the different steps of the model building

	AIC	Parameters
Multinomial	19720	3
Prop. Odds	14448	14
Over-dispersion	14441	13
Serial	*10618*	14
Markov	13426	14
Cumulative	**12858**	13

Prop. Odds: proportional-odds with covariates.

dependence is enough to predict the outcome at the next time point without any regression model. As the model is dynamic, a wrong prediction is rapidly overcome by automatically readjusting the model before the following prediction. Hence, the underlying population mean curve is unnecessary and can simply be described by a simple intercept model.

Due to the presence of a strong dependence and the lack of variability among the responses in this particular trial, we decided that the serial update is not suitable to model the problem at hand. Among the remaining models, the cumulative model has the lowest AIC (12858).

The parameter estimates from this model, and their standard errors, are presented in Table 7.2. The infusion rate coefficient (β_4) is positive implying that an increased rate leads to higher response levels. The sum of exponentials used to induce a nonlinear regression curves requires the time covariate measured in seconds to be rescaled. This explains the relatively large values for coefficients β_2, β_3, and β_9 which involve times up to 2089 seconds. The coefficients for the 4g (β_5) and 8g (β_6) concentration groups main effects are negative but the interaction coefficients are positive. This implies that the two higher concentration groups will have intercepts slightly lower than the reference (2.5g) group. But they will also increase faster to an appropriate signal level. Finally, the dependence introduced by the cumulative update is estimated to be 0.17.

The population predicted cumulative frequencies, for combinations of concentrations (2.5g, 4g, and 8g) and infusion rates (0.5ml/min, 1ml/min, 2ml/min, and 4ml/min) which occurred, are presented in Figure 7.2. It can be seen from these plots that, as for the serial update, the long series of identical response cause a distortion. Indeed, the probabilities of being in the highest category (level 3: excess Doppler signal) are significantly increased.

A rough estimate of the time a patient would spend at an appropriate Doppler signal level, can be obtained from these mean predictions. Table 7.3 summarizes the time spent at the required level depending on the concentration of drug administered and whether the patient requires a particular adjustment of its infusion rate during the trial.

Recursive fitted response levels are presented in Figure 7.3 for three indi-

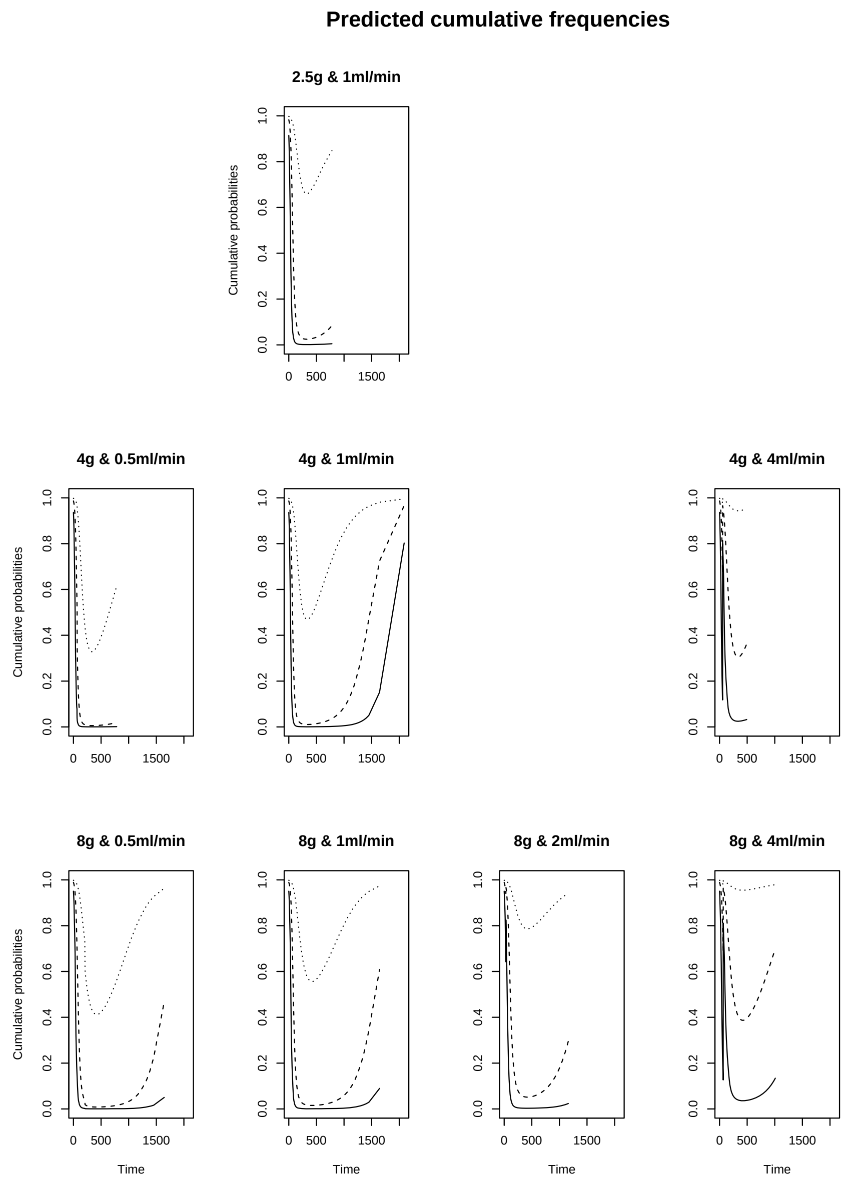

Fig. 7.2. Mean predicted cumulative frequencies for the 3 drug concentrations (2.5g, 4g, and 8g), and by the 4 infusion rate (0.5ml/min, 1ml/min, 2ml/min, and 4ml/min).

Table 7.2. Estimates for the non-linear proportional-odds model with cumulative dependence.

	Estimates	Standard Errors
Main effects (Intercepts)		
Level 1	-12.88	0.161
Level 2	-10.00	0.149
Level 3	-5.50	0.152
Regression		
β_1	3.94	0.059
β_2	87.55	1.363
β_3	1626.00	31.331
β_4	1.25	0.054
β_5	-0.90	0.070
β_6	-1.56	0.132
β_7	0.51	0.062
β_8	0.91	0.049
β_9	443.87	51.003
Dependence		
δ	0.17	—

Table 7.3. Estimated time spent at an appropriate or excess Doppler signal level.

	Infusion rates			
Conc.	0.5ml/min	1ml/min	2ml/min	4ml/min
2.5g	—	11 min, 46 sec	—	—
4.0g	11 min, 57 sec*	23 min, 06 sec**	—	5 min, 17 sec
8.0g	25 min, 59 sec*	22 min, 49 sec	17 min, 31 sec	8 min, 24 sec

*: includes approximately 7 minutes and 30 seconds spent at an excess signal level.

**: includes approximately 3 minutes spent at an excess signal level.

viduals in each concentration group who remained at an infusion rate of 1ml/min during the entire trial. The computed recursive means yield smooth curves typical of a cumulative dependence. The dynamic process can clearly be observed from these curves. Indeed, all recursive mean curves follow the underlying population curve but are slightly pulled towards a patient's observations. The recursive highest probabilities can be very useful as they represent a specific response category. These curves follow each patient's observations reasonable well.

7.2 The rhinitis allergy clinical trial

The aim of this seasonal rhinitis clinical trial was to relieve patients from six different types of symptoms due to rhinitis. The 416 patients enrolled were random-

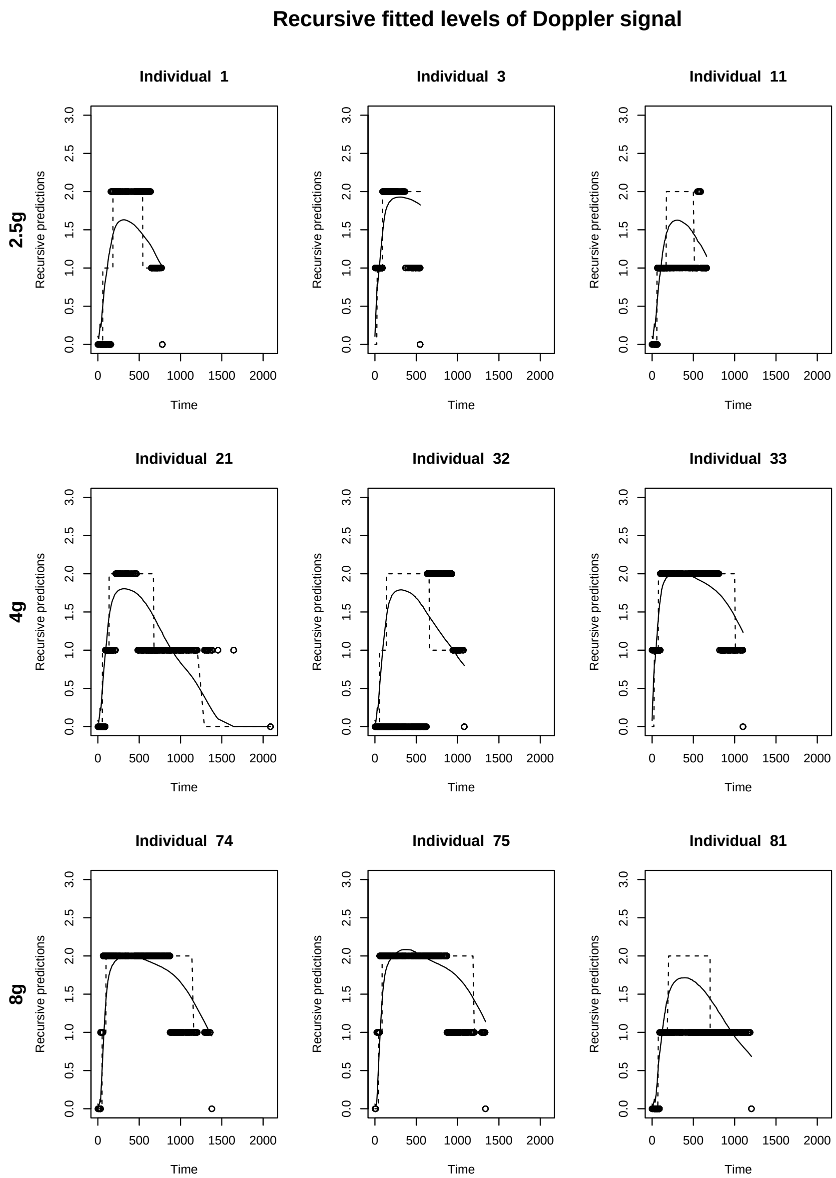

Fig. 7.3. Recursive fitted response levels for the 3 treatment groups (2.5g, 4g, and 8g). Solid line: mean fitted response levels. Dashed line: highest probability fitted response levels. Circles: observations.

ized into three groups: treatment A, treatment B, and placebo. The six symptoms recorded are blockage on waking, blockage during the day, sneezing, nasal itching, runny nose, and eye watering. Each of these were recorded by the patients themselves on a 0 to 3 ordered scale for 28 consecutive days. The presence of the symptom respectively increases with the scale (no symptom, little, moderate, bad). Further details of the trial can be found in Lindsey *et al.* (1997).

Due to missing observations, blockage on waking has a total of 10605 observations, sneezing has a total of 10648 observations, and the other symptoms have a total of 10647 observations. These missing observations occur among the first two treatment days, patients dropping out half way through, one patient who was not included due to unknown treatment, and another patient who withdrew before treatment assignment. Because the missingness does not appear to depend on the treatment, it will be assumed to be missing at random. This implies that the data are not balanced because each patient has a response series of different length and possibly different recording times.

7.3 The tick activity data

This dataset consists of 28 ticks from the same species observed over a period of 181 days. Colored numbered tags, used by bee-keepers to identify the queens, were applied to the ticks for individual identification. They are then placed into 4 different columns corresponding to their originating locality (all in Zambia, Africa): Michembo, Genda, Nkolowondo, and Lundazi. The activity of each tick is measured daily and recorded into three categories: active, not active, not visible. Several environmental variables are also available such as relative humidity, temperature, and VDP (a nonlinear combination of relative humidity and temperature.

7.4 Three Mile Island Stress data

This study investigates the psychological effect due to the nuclear accident at the Three Mile Island power plant in the spring of 1979. This is measured by the changes in levels of stress of mothers of young children living within 10 miles of the plant. Four interviews were conducted following the accident: winter 1979, spring 1980, fall 1981, and fall 1982. At each interview, subjects are classified into one of three categories, low, medium, or high stress, based on a composite score from a 90-item checklist. Observations are available for the 267 subjects who completed all four interviews. These are stratified into two groups, those living within 5 miles and 10 miles from the plant. Further details of this study can be found in Conaway (1989).

Appendix A
Linear categorical regression

A.1 Global notation

- Number of categories: `nc`
- Number of covariates: `ncv`
- Frequency for response category *k* at line *i* of frequency table: n_{ik}
- Regression intercept coefficient: α
- Regression intercept coefficient for response category *k*: α_k
- Regression covariate coefficient for covariate *l*: β_l
- Regression covariate coefficient for response category *k* of covariate *l*: β_{kl}
- Covariate *l* at line *i* of frequency table: x_{il}

A.2 Binary

$$
\begin{aligned}
a_i &= \exp\left(\alpha + \sum_{l=1}^{\texttt{ncv}} \beta_l x_{il}\right) \\
p_i &= \frac{a_i}{1+a_i} \\
q_i &= \frac{1}{1+a_i} = 1 - p_i
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\mathbf{1} \le \boldsymbol{r}, \boldsymbol{s} \le \mathbf{ncv}}$$

$$
\begin{aligned}
\mathrm{L} &= \prod_i \left[(p_i)^{n_{i1}} (q_i)^{n_{i\bullet} - n_{i1}}\right] \\
\mathrm{ll} &= \sum_i \left[n_{i1} \left(\alpha + \sum_{l=1}^{\texttt{ncv}} \beta_l x_{il}\right) - n_{i\bullet} \ln(1+a_i)\right] \\
\mathrm{S}_\alpha &= n_{\bullet 1} - \sum_i [n_{i\bullet} p_i] \\
\mathrm{S}_{\beta_r} &= \sum_i [n_{i1} x_{ir}] - \sum_i [n_{i\bullet} x_{ir} p_i]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{I}_{\alpha\alpha} &= -\sum_i [n_{i\bullet} p_i q_i] \\
\mathrm{I}_{\alpha\beta_r} &= -\sum_i [n_{i\bullet} x_{ir} p_i q_i] \\
&= \mathrm{I}_{\beta_r\alpha} \\
\mathrm{I}_{\beta_r\beta_s} &= -\sum_i [n_{i\bullet} x_{ir} x_{is} p_i q_i] \\
&= \mathrm{I}_{\beta_s\beta_r}
\end{aligned}
$$

A.3 Multinomial

$$\boxed{\mathbf{1} \le \boldsymbol{k} \le (\mathbf{nc}-\mathbf{1})}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_{kl} x_{il}\right) \\
p_{ik} &= \frac{a_{ik}}{1 + \sum_{h=1}^{\mathtt{nc}-1} a_{ih}} \\
q_i &= \frac{1}{1 + \sum_{h=1}^{\mathtt{nc}-1} a_{ih}} = 1 - \sum_{h=1}^{\mathtt{nc}-1} p_{ih}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l} \mathbf{1} \le \boldsymbol{u}, \boldsymbol{v} \le (\mathbf{nc}-\mathbf{1}) \\ \boldsymbol{u} \ne \boldsymbol{v} \\ \mathbf{1} \le \boldsymbol{r}, \boldsymbol{s} \le \mathbf{ncv} \end{array}}$$

$$
\begin{aligned}
\mathrm{L} &= \prod_i \left[(q_i)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \prod_{h=1}^{\mathtt{nc}-1} (p_{ih})^{n_{ih}} \right] \\
\mathrm{ll} &= \sum_i \left[\sum_{h=1}^{\mathtt{nc}-1} \left\{ n_{ih} \left(\alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_{hl} x_{il} \right) \right\} - n_{i\bullet} \ln \left(1 + \sum_{h=1}^{\mathtt{nc}-1} a_{ih} \right) \right] \\
\mathrm{S}_{\alpha_u} &= n_{\bullet u} - \sum_i [n_{i\bullet} p_{iu}] \\
\mathrm{S}_{\beta_{ur}} &= \sum_i [n_{iu} x_{ir}] - \sum_i [n_{i\bullet} x_{ir} p_{iu}]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{I}_{\alpha_u\alpha_u} &= -\sum_i [n_{i\bullet} p_{iu}(1-p_{iu})] \\
\mathrm{I}_{\alpha_u\alpha_v} &= \sum_i [n_{i\bullet} p_{iu} p_{iv}] \\
&= \mathrm{I}_{\alpha_v\alpha_u} \\
\mathrm{I}_{\alpha_u\beta_{ur}} &= -\sum_i [n_{i\bullet} x_{ir} p_{iu}(1-p_{iu})] \\
&= \mathrm{I}_{\beta_{ur}\alpha_u} \\
\mathrm{I}_{\alpha_u\beta_{vr}} &= \sum_i [n_{i\bullet} x_{ir} p_{iu} p_{iv}] \\
&= \mathrm{I}_{\beta_{vr}\alpha_u} \\
\mathrm{I}_{\beta_{ur}\beta_{us}} &= -\sum_i [n_{i\bullet} x_{ir} x_{is} p_{iu}(1-p_{iu})] \\
&= \mathrm{I}_{\beta_{im}\beta_{is}} \\
\mathrm{I}_{\beta_{ur}\beta_{vs}} &= \sum_i [n_{i\bullet} x_{ir} x_{is} p_{iu} p_{iv}] \\
&= \mathrm{I}_{\beta_{vs}\beta_{ur}}
\end{aligned}
$$

A.4 Simplified multinomial

$$\boxed{\mathbf{1 \le \boldsymbol{k} \le (nc-1)}}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il}\right) \\
p_{ik} &= \frac{a_{ik}}{1+\sum_{h=1}^{\mathtt{nc}-1} a_{ih}} \\
q_i &= \frac{1}{1+\sum_{h=1}^{\mathtt{nc}-1} a_{ih}} = 1 - \sum_{h=1}^{\mathtt{nc}-1} p_{ih}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l}\mathbf{1 \le \boldsymbol{u}, \boldsymbol{v} \le (nc-1)} \\ \boldsymbol{u \ne v} \\ \mathbf{1 \le \boldsymbol{r}, \boldsymbol{s} \le ncv}\end{array}}$$

$$
\begin{aligned}
\mathrm{L} &= \prod_i \left[(q_i)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \prod_{h=1}^{\mathtt{nc}-1} (p_{ih})^{n_{ih}} \right] \\
\mathrm{ll} &= \sum_i \left[\sum_{h=1}^{\mathtt{nc}-1} \left\{ n_{ih} \left(\alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \right) \right\} - n_{i\bullet} \ln \left(1 + \sum_{h=1}^{\mathtt{nc}-1} a_{ih} \right) \right]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{S}_{\alpha_u} &= n_{\bullet u} - \sum_i [n_{i\bullet} p_{iu}] \\
\mathrm{S}_{\beta_r} &= \sum_i \left[x_{ir} \sum_{h=1}^{\mathtt{nc}-1} n_{ih} \right] - \sum_i [n_{i\bullet} x_{ir} (1-q_i)] \\
\mathrm{I}_{\alpha_u \alpha_u} &= -\sum_i [n_{i\bullet} p_{iu} (1-p_{iu})] \\
\mathrm{I}_{\alpha_u \alpha_v} &= \sum_i [n_{i\bullet} p_{iu} p_{iv}] \\
&= \mathrm{I}_{\alpha_v \alpha_u} \\
\mathrm{I}_{\alpha_u \beta_r} &= -\sum_i [n_{i\bullet} x_{ir} p_{iu} q_i] \\
&= \mathrm{I}_{\beta_r \alpha_u} \\
\mathrm{I}_{\beta_r \beta_s} &= -\sum_i [n_{i\bullet} x_{ir} x_{is} q_i (1-q_i)] \\
&= \mathrm{I}_{\beta_s \beta_r}
\end{aligned}
$$

A.5 Continuation ratio

As categories are ordered, only the first or the last level can be set as the reference category. This yields the following two distinct cases.

UPWARDS:

$$\boxed{\mathbf{1 \leq \boldsymbol{k} \leq (nc-1)}}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \right) \\
p_{ik} &= \frac{a_{ik}}{1+a_{ik}} \\
q_{ik} &= \frac{1}{1+a_{ik}} = 1 - p_{ik}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l} \mathbf{1 \leq \boldsymbol{u}, \boldsymbol{v} \leq (nc-1)} \\ \boldsymbol{u \neq v} \\ \mathbf{1 \leq \boldsymbol{r}, \boldsymbol{s} \leq ncv} \end{array}}$$

$$
\begin{aligned}
\mathrm{L} &= \prod_i \left[(p_{i1})^{n_{i1}} (q_{i1})^{n_{i2}} \dots (p_{i,\mathtt{nc}-1})^{\sum_{h=1}^{\mathtt{nc}-1} n_{ih}} (q_{i,\mathtt{nc}-1})^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \right] \\
\mathrm{ll} &= \sum_i \left[\sum_{h=1}^{\mathtt{nc}-1} \left\{ \left(\alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \right) \sum_{g=1}^{h} n_{ig} - \ln(1+a_{ih}) \sum_{g=1}^{h+1} n_{ig} \right\} \right]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{S}_{\alpha_u} &= \sum_{g=1}^{u} n_{\bullet g} - \sum_i \left[\sum_{g=1}^{u}\{n_{ig}p_{iu}\}\right] - \sum_i [n_{i,u+1}p_{iu}] \\
\mathrm{S}_{\beta_r} &= \sum_i \left[x_{ir}\sum_{g=1}^{\mathtt{nc}-1}\{(\mathtt{nc}-g)n_{ig}\}\right] \\
&\qquad - \sum_i \left[x_{ir}\sum_{h=1}^{\mathtt{nc}-1}\left(n_{ih}\sum_{g=h}^{\mathtt{nc}-1}p_{ig}\right)\right] - \sum_i \left[x_{ir}\sum_{h=1}^{\mathtt{nc}-1}(n_{i,h+1}p_{ih})\right] \\
\mathrm{I}_{\alpha_u\alpha_u} &= -\sum_i \left[p_{iu}q_{iu}\sum_{g=1}^{u} n_{ig}\right] - \sum_i [n_{i,u+1}p_{iu}q_{iu}] \\
\mathrm{I}_{\alpha_u\alpha_v} &= 0 \\
&= \mathrm{I}_{\alpha_v\alpha_u} \\
\mathrm{I}_{\alpha_u\beta_r} &= -\sum_i \left[x_{ir}p_{iu}q_{iu}\sum_{g=1}^{u} n_{ig}\right] - \sum_i [n_{i,u+1}x_{ir}p_{iu}q_{iu}] \\
&= \mathrm{I}_{\beta_r\alpha_u} \\
\mathrm{I}_{\beta_r\beta_s} &= -\sum_i \left[x_{ir}x_{is}\sum_{h=1}^{\mathtt{nc}-1}\left(n_{ih}\sum_{g=h}^{\mathtt{nc}-1}p_{ig}q_{ig}\right)\right] \\
&\qquad - \sum_i \left[x_{ir}x_{is}\sum_{h=1}^{\mathtt{nc}-1}(n_{i,h+1}p_{ih}q_{ih})\right] \\
&= \mathrm{I}_{\beta_s\beta_r}
\end{aligned}
$$

DOWNWARDS:

$$\boxed{\mathbf{2 \le \boldsymbol{k} \le nc}}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(\alpha_k + \sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}\right) \\
p_{ik} &= \frac{a_{ik}}{1+a_{ik}} \\
q_{ik} &= \frac{1}{1+a_{ik}} = 1 - p_{ik}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l} \mathbf{2 \le \boldsymbol{u}, \boldsymbol{v} \le nc} \\ \boldsymbol{u \ne v} \\ \mathbf{1 \le \boldsymbol{r}, \boldsymbol{s} \le ncv} \end{array}}$$

$$\mathrm{L} = \prod_i \left[(q_{i2})^{n_{i\bullet}-\sum_{h=2}^{\mathtt{nc}} n_{ih}}(p_{i2})^{\sum_{h=2}^{\mathtt{nc}} n_{ih}} \ldots (q_{i,\mathtt{nc}})^{n_{i,\mathtt{nc}-1}}(p_{i,\mathtt{nc}})^{n_{i,\mathtt{nc}}}\right]$$

$$
\begin{aligned}
\mathrm{ll} &= \sum_i \left[\sum_{h=2}^{\mathtt{nc}} \left\{ \left(\alpha_h + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \right) \sum_{g=h}^{\mathtt{nc}} n_{ig} - \ln(1+a_{ih}) \sum_{g=h-1}^{\mathtt{nc}} n_{ig} \right\} \right] \\
\mathrm{S}_{\alpha_u} &= \sum_{g=u}^{\mathtt{nc}} n_{\bullet g} - \sum_i \left[\sum_{g=u}^{\mathtt{nc}} \{ n_{ig} p_{iu} \} \right] - \sum_i [n_{i,u-1} p_{iu}] \\
\mathrm{S}_{\beta_r} &= \sum_i \left[x_{ir} \sum_{h=2}^{\mathtt{nc}} \sum_{g=h}^{\mathtt{nc}} n_{ig} \right] - \sum_i \left[x_{ir} \sum_{h=2}^{\mathtt{nc}} \left(n_{ih} \sum_{g=2}^{h} p_{ig} \right) \right] \\
&\qquad - \sum_i \left[x_{ir} \sum_{h=2}^{\mathtt{nc}} (n_{i,h-1} p_{ih}) \right] \\
\mathrm{I}_{\alpha_u \alpha_u} &= -\sum_i \left[p_{iu} q_{iu} \sum_{g=u}^{\mathtt{nc}} n_{ig} \right] - \sum_i [p_{iu} q_{iu} n_{i,u-1}] \\
\mathrm{I}_{\alpha_u \alpha_v} &= 0 \\
&= \mathrm{I}_{\alpha_v \alpha_u} \\
\mathrm{I}_{\alpha_u \beta_r} &= -\sum_i \left[x_{ir} p_{iu} q_{iu} \sum_{g=u}^{\mathtt{nc}} n_{ig} \right] - \sum_i [x_{ir} p_{iu} q_{iu} n_{i,u-1}] \\
&= \mathrm{I}_{\beta_r \alpha_u} \\
\mathrm{I}_{\beta_r \beta_s} &= -\sum_i \left[x_{ir} x_{is} \sum_{h=2}^{\mathtt{nc}} \left(n_{ih} \sum_{g=2}^{h} p_{ig} q_{ig} \right) \right] - \sum_i \left[x_{ir} x_{is} \sum_{h=2}^{\mathtt{nc}} (n_{i,h-1} p_{ih} q_{ih}) \right] \\
&= \mathrm{I}_{\beta_s \beta_r}
\end{aligned}
$$

A.6 Proportional odds

$$\boxed{\mathbf{1} \le \boldsymbol{k} \le \mathbf{(nc{-}1)}}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(\alpha_k + \sum_{l=1}^{\mathtt{ncv}} \beta_l x_{il} \right) \\
p_{ik} &= \frac{a_{ik}}{1+a_{ik}} \\
q_{ik} &= \frac{1}{1+a_{ik}} = 1 - p_{ik}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l} \mathbf{2 \le u \le (nc-2)} \\ \mathbf{2 \le v \le (nc-3)} \\ \mathbf{1 \le w \le (nc-3)} \\ \mathbf{(w+2) \le z \le (nc-1)} \\ \mathbf{1 \le r, s \le ncv} \end{array}}$$

$$\mathrm{L} = \prod_i [(p_{i1})^{n_{i1}}(p_{i2}-p_{i1})^{n_{i2}}\dots(q_{i,\mathtt{nc}-1})^{n_{i,\mathtt{nc}}}]$$

$$\begin{aligned} \mathrm{ll} = \sum_i \Bigg[& n_{i1}\left(\alpha_1 + \sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}\right) - n_{i1}\ln(1+a_{i1}) \\ & + \sum_{g=2}^{\mathtt{nc}-1}\{n_{ig}\ln(p_{ig}-p_{i,g-1})\} - n_{i,\mathtt{nc}}\ln(1+a_{i,\mathtt{nc}-1})\Bigg] \end{aligned}$$

$$\begin{aligned}
\mathrm{S}_{\alpha_1} &= n_{\bullet 1} - \sum_i [n_{i1}p_{i1}] - \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(\sum_i\left[n_{i2}\frac{p_{i1}q_{i1}}{p_{i2}-p_{i1}}\right]\right) \\
\mathrm{S}_{\alpha_u} &= \underset{\mathtt{nc}>3}{\mathrm{IF}}\left(\sum_i\left[n_{iu}\frac{p_{iu}q_{iu}}{p_{iu}-p_{i,u-1}}\right] - \sum_i\left[n_{i,u+1}\frac{p_{iu}q_{iu}}{p_{i,u+1}-p_{iu}}\right]\right) \\
\mathrm{S}_{\alpha_{\mathtt{nc}-1}} &= \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(\sum_i\left[n_{i,\mathtt{nc}-1}\frac{p_{i,\mathtt{nc}-1}q_{i,\mathtt{nc}-1}}{p_{i,\mathtt{nc}-1}-p_{i,\mathtt{nc}-2}}\right]\right) - \sum_i[n_{i,\mathtt{nc}}p_{i,\mathtt{nc}-1}] \\
\mathrm{S}_{\beta_r} &= \sum_i\sum_{h=1}^{\mathtt{nc}-1}[n_{ih}x_{ir}q_{ih}] - \sum_i\sum_{h=2}^{\mathtt{nc}}[n_{ih}x_{ir}p_{i,h-1}] \\
\mathrm{I}_{\alpha_1\alpha_1} &= -\sum_i[n_{i1}p_{i1}q_{i1}] - \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(\sum_i\left[n_{i2}p_{i1}q_{i1}\left\{1+\frac{p_{i2}q_{i2}}{(p_{i2}-p_{i1})^2}\right\}\right]\right) \\
\mathrm{I}_{\alpha_u\alpha_u} &= \underset{\mathtt{nc}>3}{\mathrm{IF}}\left(-\sum_i\left[n_{iu}p_{iu}q_{iu}\left\{1+\frac{p_{i,u-1}q_{i,u-1}}{(p_{iu}-p_{i,u-1})^2}\right\}\right]\right. \\
&\qquad\left. -\sum_i\left[n_{i,u+1}p_{iu}q_{iu}\left\{1+\frac{p_{i,u+1}q_{i,u+1}}{(p_{i,u+1}-p_{iu})^2}\right\}\right]\right) \\
\mathrm{I}_{\alpha_{\mathtt{nc}-1}\alpha_{\mathtt{nc}-1}} &= \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(-\sum_i\left[n_{i,\mathtt{nc}-1}p_{i,\mathtt{nc}-1}q_{i,\mathtt{nc}-1}\left\{1+\frac{p_{i,\mathtt{nc}-2}q_{i,\mathtt{nc}-2}}{(p_{i,\mathtt{nc}-1}-p_{i,\mathtt{nc}-2})^2}\right\}\right]\right) \\
\mathrm{I}_{\alpha_1\alpha_2} &= \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(\sum_i\left[n_{i2}p_{i1}q_{i1}\frac{p_{i2}q_{i2}}{(p_{i2}-p_{i1})^2}\right]\right) \\
&= \mathrm{I}_{\alpha_2\alpha_1}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{I}_{\alpha_v\alpha_{v+1}} &= \underset{\mathtt{nc}>4}{\mathrm{IF}}\left(\sum_i\left[n_{i,v+1}p_{iv}q_{iv}\frac{p_{i,v+1}q_{i,v+1}}{(p_{i,v+1}-p_{iv})^2}\right]\right)\\
&= \mathrm{I}_{\alpha_{v+1}\alpha_v}\\
\mathrm{I}_{\alpha_{\mathtt{nc}-2}\alpha_{\mathtt{nc}-1}} &= \underset{\mathtt{nc}>3}{\mathrm{IF}}\left(\sum_i\left[n_{i,\mathtt{nc}-1}p_{i,\mathtt{nc}-1}q_{i,\mathtt{nc}-1}\frac{p_{i,\mathtt{nc}-2}q_{i,\mathtt{nc}-2}}{(p_{i,\mathtt{nc}-1}-p_{i,\mathtt{nc}-2})^2}\right]\right)\\
&= \mathrm{I}_{\alpha_{\mathtt{nc}-1}\alpha_{\mathtt{nc}-2}}\\
\mathrm{I}_{\alpha_w\alpha_z} &= \underset{\mathtt{nc}>3}{\mathrm{IF}}\ 0\\
&= \mathrm{I}_{\alpha_z\alpha_w}\\
\mathrm{I}_{\alpha_1\beta_r} &= -\sum_i[n_{i1}x_{ir}p_{i1}q_{i1}] - \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(\sum_i[n_{i2}x_{ir}p_{i1}q_{i1}]\right)\\
&= \mathrm{I}_{\beta_r\alpha_1}\\
\mathrm{I}_{\alpha_u\beta_r} &= \underset{\mathtt{nc}>u}{\mathrm{IF}}\left(-\sum_i[n_{iu}x_{ir}p_{iu}q_{iu}]\right) - \underset{\mathtt{nc}>(u+1)}{\mathrm{IF}}\left(\sum_i[n_{i,u+1}x_{ir}p_{iu}q_{iu}]\right)\\
&= \mathrm{I}_{\beta_r\alpha_u}\\
\mathrm{I}_{\alpha_{\mathtt{nc}-1}\beta_r} &= \underset{\mathtt{nc}>2}{\mathrm{IF}}\left(-\sum_i[n_{i,\mathtt{nc}-1}x_{ir}p_{i,\mathtt{nc}-1}q_{i,\mathtt{nc}-1}]\right)\\
&\qquad\qquad -\sum_i[n_{i,\mathtt{nc}}x_{ir}p_{i,\mathtt{nc}-1}q_{i,\mathtt{nc}-1}]\\
&= \mathrm{I}_{\beta_r\alpha_{\mathtt{nc}-1}}\\
\mathrm{I}_{\beta_r\beta_s} &= -\sum_i\sum_{h=1}^{\mathtt{nc}-1}[n_{ih}x_{ir}x_{is}p_{ih}q_{ih}] - \sum_i\sum_{h=2}^{\mathtt{nc}}[n_{ih}x_{ir}x_{is}p_{i,h-1}q_{i,h-1}]\\
&= \mathrm{I}_{\beta_s\beta_r}
\end{aligned}
$$

A.7 Adjacent category

$$\boxed{\mathbf{1 \le \mathit{k} \le (nc-1)}}$$

$$
\begin{aligned}
a_{ik} &= \exp\left(-\sum_{g=1}^{k}\alpha_g + \sum_{l=1}^{\mathtt{ncv}}\beta_l x_{il}\right)\\
p_{ik} &= \frac{a_{ik}}{1+\sum_{h=1}^{\mathtt{nc}-1}a_{ih}}\\
q_i &= \frac{1}{1+\sum_{h=1}^{\mathtt{nc}-1}a_{ih}} = 1-\sum_{h=1}^{\mathtt{nc}-1}p_{ih}
\end{aligned}
$$

LIKELIHOOD, SCORE EQUATIONS, AND INFORMATION MATRIX

$$\boxed{\begin{array}{l}\mathbf{1 \le \boldsymbol{u}, \boldsymbol{v} \le (nc-1)}\\ \mathbf{1 \le \boldsymbol{r}, \boldsymbol{s} \le ncv}\end{array}}$$

$$
\begin{aligned}
\mathrm{L} &= \prod_i \left[(q_i)^{n_{i\bullet} - \sum_{h=1}^{\mathtt{nc}-1} n_{ih}} \prod_{h=1}^{\mathtt{nc}-1} (p_{ih})^{n_{ih}} \right] \\
\mathrm{ll} &= -\sum_i \left[\sum_{h=1}^{\mathtt{nc}-1} \left\{ n_{ih} \left(h \sum_{l=1}^{\mathtt{ncv}} (\beta_l x_{il}) + \sum_{g=1}^{h} \alpha_g \right) \right\} + n_{i\bullet} \ln \left(1 + \sum_{h=1}^{\mathtt{nc}-1} a_{ih} \right) \right] \\
\mathrm{S}_{\alpha_u} &= -\sum_{g=u}^{\mathtt{nc}-1} n_{\bullet g} + \sum_i \left[n_{i\bullet} \sum_{g=u}^{\mathtt{nc}-1} p_{ig} \right] \\
\mathrm{S}_{\beta_r} &= -\sum_i \left[x_{ir} \sum_{h=1}^{\mathtt{nc}-1} (h n_{ih}) \right] + \sum_i \left[n_{i\bullet} x_{ir} \sum_{h=1}^{\mathtt{nc}-1} (h p_{ih}) \right] \\
\mathrm{I}_{\alpha_u \alpha_v} &= -\sum_i \left[n_{i\bullet} \left(\sum_{g=v}^{\mathtt{nc}-1} p_{ig} \right) \left(1 - \sum_{g=u}^{\mathtt{nc}-1} p_{ig} \right) \right] \\
&= \mathrm{I}_{\alpha_v \alpha_u} \\
\mathrm{I}_{\alpha_u \beta_r} &= -\sum_i \left[n_{i\bullet} x_{ir} \sum_{g=u}^{\mathtt{nc}-1} \left\{ g p_{ig} - p_{ig} \sum_{h=1}^{\mathtt{nc}-1} (h p_{ih}) \right\} \right] \\
&= \mathrm{I}_{\beta_r \alpha_u} \\
\mathrm{I}_{\beta_r \beta_s} &= -\sum_i \left[n_{i\bullet} x_{ir} x_{is} \sum_{h=1}^{\mathtt{nc}-1} \left\{ h^2 p_{ih} - h p_{ih} \sum_{g=1}^{\mathtt{nc}-1} (g p_{ig}) \right\} \right] \\
&= \mathrm{I}_{\beta_s \beta_r}
\end{aligned}
$$

Appendix B
Distributions

This appendix defines the parametrization of distributions that has been used in most cases. The first section gives the relationships among the cdf, survival function, density, integrated hazard, and hazard in the univariate case. Then, various mathematical functions such as the gamma, beta, Bessel, and standard Gaussian function are defined in Section B.2. Univariate distributions are covered in Sections B.3 and B.4. Multivariate distributions are covered in Section B.5. As can be noticed in these sections, there is no unique multivariate extension from a univariate distribution. Indeed, infinitely many multivariate distributions can have the same univariate marginal distributions.

B.1 General univariate relationships

The standard relationships linking the cdf or survival function to the density, integrated hazard, and hazard of a same univariate distribution are

- Density $f(y) = \frac{\partial F(y)}{\partial y}$
- CDF $F(y) = 1 - S(y)$
- Survival function $S(y) = 1 - F(y)$
- Integrated hazard $H(y) = -\ln[S(y)]$
- Hazard $h(y) = \frac{f(y)}{S(y)}$

Note that most distributions are defined from their density and do not have a closed form cdf and survival function. Nevertheless, the few distributions that on the contrary have explicit cdfs and survival functions will be of major interest because the responses under study are nominal or ordinal. Note that if the cdf and survival function have a closed form and their first derivative exists then all other relationships can also be explicitly worked out.

B.2 Functions in distributions

GAMMA FUNCTION

The gamma function is

$$\Gamma(a) = \int_0^\infty z^{\alpha-1} \mathrm{e}^{-z} \mathrm{d}z \tag{B.1}$$

where the parameter α is positive. Note that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.
(Johnson and Kotz, 1969, p. 6)

INCOMPLETE GAMMA FUNCTION

The incomplete gamma function is

$$\Gamma(y, a) = \int_0^y z^{a-1} \mathrm{e}^{-z} \mathrm{d}z = \Gamma(a) F(y)$$

where $F(y)$ is a one parameter gamma cdf and the parameter a is positive.
(Johnson and Kotz, 1969, p. 6; Johnson and Kotz, 1970a, p. 167)

FRACTIONAL BESSEL FUNCTION OF THE THIRD KIND

The fractional Bessel function of the third kind is

$$K\left(a, |b|\right) = \frac{1}{2} \int_0^\infty z^{b-1} \mathrm{e}^{-\frac{a\left(z+\frac{1}{z}\right)}{2}} \mathrm{d}z \tag{B.2}$$

and this has a closed form when $b = -\frac{1}{2}$.

$$K\left(a, \frac{1}{2}\right) = \sqrt{\frac{\pi}{2a}} \mathrm{e}^{-a}$$

(Jørgensen, 1982, p. 170; Seshadri, 1993, p. 46; Hougaard, 2000, pp. 510–511)

BETA FUNCTION

The beta function is

$$B(\alpha, \beta) = \int_0^\infty z^{\alpha-1} (1-z)^{\beta-1} \mathrm{d}z = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{B.3}$$

where the parameters α and $beta$ are positive.
(Johnson and Kotz, 1970b, pp. 37–38 and 40)

STANDARDIZED GAUSSIAN FUNCTION

The standardized Gaussian function is

$$\Phi(z) = \frac{\int_{-\infty}^{y} \mathrm{e}^{-\frac{z^2}{2}} \mathrm{d}z}{\sqrt{2\pi}} \tag{B.4}$$

(Johnson and Kotz, 1970a, p. 40)

B.3 Continuous univariate distributions

EXPONENTIAL DISTRIBUTION

Two parameter distribution

- Density $f(y) = \frac{e^{-\frac{y-\mu}{\sigma}}}{\sigma} = \lambda e^{-\lambda(y-\mu)}$
- CDF $F(y) = 1 - e^{-\lambda(y-\mu)}$
- Survival function $S(y) = e^{-\lambda(y-\mu)}$
- Integrated hazard $H(y) = \lambda(y-\mu)$
- Hazard $h(y) = \lambda$

The responses y are non-negative and $\lambda = \frac{1}{\sigma}$. The location is $\mu < y$, and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970a, p. 207)

One parameter distributions

- Density $f(y) = \frac{e^{-\frac{y}{\sigma}}}{\sigma} = \lambda e^{-\lambda y}$ (B.5)
- CDF $F(y) = 1 - e^{-\lambda y}$
- Survival function $S(y) = e^{-\lambda y}$ (B.6)
- Integrated hazard $H(y) = \lambda y$
- Hazard $h(y) = \lambda$

The responses y are non-negative and $\lambda = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$) and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970a, p. 207)

- Density $f(y) = e^{-(y-\mu)}$
- CDF $F(y) = 1 - e^{-(y-\mu)}$
- Survival function $S(y) = e^{-(y-\mu)}$
- Integrated hazard $H(y) = y - \mu$
- Hazard $h(y) = 1$

The responses y are non-negative. The location is $\mu < y$ and the scale is set to unity $\left(\lambda = \frac{1}{\sigma} = 1\right)$.
(Johnson and Kotz, 1970a, p. 207)

Zero parameter distribution

$$
\begin{aligned}
&\bullet \text{ Density} && f(y) = \mathrm{e}^{-y} \\
&\bullet \text{ CDF} && F(y) = 1-\mathrm{e}^{-y} \\
&\bullet \text{ Survival function} && S(y) = \mathrm{e}^{-y} \\
&\bullet \text{ Integrated hazard} && H(y) = y \\
&\bullet \text{ Hazard} && h(y) = 1
\end{aligned}
\tag{B.7}
$$

The responses y are non-negative. The location is set to zero ($\mu = 0$) and the scale is set to unity $\left(\lambda = \frac{1}{\sigma} = 1\right)$. This is the standard form called the unit exponential. (Johnson and Kotz, 1970a, p. 207)

GAMMA DISTRIBUTION

Three parameter distribution

$$
\begin{aligned}
&\bullet \text{ Density} && f(y) = \frac{(y-\mu)^{\alpha-1}\mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\sigma^{\alpha}\Gamma(\alpha)} \\
& && = \frac{\beta^{\alpha}(y-\mu)^{\alpha-1}\mathrm{e}^{-\beta(y-\mu)}}{\Gamma(\alpha)} \\
&\bullet \text{ CDF} && F(y) = \frac{\Gamma(\beta(y-\mu),\alpha)}{\Gamma(\alpha)} \\
&\bullet \text{ Survival function} && S(y) = \frac{\Gamma(\alpha)-\Gamma(\beta(y-\mu),\alpha)}{\Gamma(\alpha)} \\
&\bullet \text{ Integrated hazard} && H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(\beta(y-\mu),\alpha)] \\
&\bullet \text{ Hazard} && h(y) = \frac{y^{\alpha-1}\mathrm{e}^{-y}}{\Gamma(\alpha)-\Gamma(\beta(y-\mu),\alpha)}
\end{aligned}
$$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is $\mu < y$, the scale is $\sigma > 0$, and the shape is $\alpha > 0$. The exponential distribution is obtained when $\alpha = 1$. (Johnson and Kotz, 1970a, p. 166)

Two parameter distributions

$$
\begin{aligned}
&\bullet \text{ Density} && f(y) = \frac{y^{\alpha-1}\mathrm{e}^{-\frac{y}{\sigma}}}{\sigma^{\alpha}\Gamma(\alpha)} \\
& && = \frac{\beta^{\alpha}y^{\alpha-1}\mathrm{e}^{-\beta y}}{\Gamma(\alpha)} \\
&\bullet \text{ CDF} && F(y) = \frac{\Gamma(\beta y,\alpha)}{\Gamma(\alpha)} \\
&\bullet \text{ Survival function} && S(y) = \frac{\Gamma(\alpha)-\Gamma(\beta y,\alpha)}{\Gamma(\alpha)} \\
&\bullet \text{ Integrated hazard} && H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(\beta y,\alpha)] \\
&\bullet \text{ Hazard} && h(y) = \frac{\beta^{\alpha}y^{\alpha-1}\mathrm{e}^{-\beta y}}{\Gamma(\alpha)-\Gamma(\beta y,\alpha)}
\end{aligned}
\tag{B.8}
$$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$), the

scale is $\sigma > 0$, and the shape is $\alpha > 0$. The theoretical mean and variance are respectively $\mathrm{E}(X) = \frac{\alpha}{\beta}$ and $\mathrm{Var}(X) = \frac{\alpha}{\beta^2}$.
(Johnson and Kotz, 1970a, p. 166)

- Density $\quad f(y) = \frac{(y-\mu)^{\alpha-1}\mathrm{e}^{-(y-\mu)}}{\Gamma(\alpha)}$
- CDF $\quad F(y) = \frac{\Gamma(y-\mu,\alpha)}{\Gamma(\alpha)}$
- Survival function $\quad S(y) = \frac{\Gamma(\alpha)-\Gamma(y-\mu,\alpha)}{\Gamma(\alpha)}$
- Integrated hazard $\quad H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(y-\mu,\alpha)]$
- Hazard $\quad h(y) = \frac{y^{\alpha-1}\mathrm{e}^{-y}}{\Gamma(\alpha)-\Gamma(y-\mu,\alpha)}$

The responses y are positive. The location is $\mu < y$, the scale is set to unity $\left(\beta = \frac{1}{\sigma} = 1\right)$, and the shape is $\alpha > 0$.
(Johnson and Kotz, 1970a, p. 166)

One parameter distribution

- Density $\quad f(y) = \frac{y^{\alpha-1}\mathrm{e}^{-y}}{\Gamma(\alpha)}$
- CDF $\quad F(y) = \frac{\Gamma(y,\alpha)}{\Gamma(\alpha)}$
- Survival function $\quad S(y) = \frac{\Gamma(\alpha)-\Gamma(y,\alpha)}{\Gamma(\alpha)}$
- Integrated hazard $\quad H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(y,\alpha)]$
- Hazard $\quad h(y) = \frac{y^{\alpha-1}\mathrm{e}^{-y}}{\Gamma(\alpha)-\Gamma(y,\alpha)}$

The responses y are positive. The location is set to zero ($\mu = 0$), the scale is set to unity $\left(\beta = \frac{1}{\sigma} = 1\right)$, and the shape is $\alpha > 0$. This is the standard form and is called an Erlang distribution when α is a positive integer.
(Johnson and Kotz, 1970a, p. 166)

GENERALIZED GAMMA DISTRIBUTION

Four parameter distribution

- Density $$f(y) = \frac{\kappa(y-\mu)^{\kappa\alpha-1}\mathrm{e}^{-\left(\frac{y-\mu}{\sigma}\right)^{\kappa}}}{\sigma^{\kappa\alpha}\Gamma(\alpha)} = \frac{\kappa(y-\mu)^{\kappa\alpha-1}(\beta^{\kappa})^{\alpha}\mathrm{e}^{-\beta^{\kappa}(y-\mu)^{\kappa}}}{\Gamma(\alpha)}$$
- CDF $$F(y) = \frac{\Gamma(\beta^{\kappa}[y-\mu]^{\kappa},\alpha)}{\Gamma(\alpha)}$$
- Survival function $$S(y) = \frac{\Gamma(\alpha)-\Gamma(\beta^{\kappa}[y-\mu]^{\kappa},\alpha)}{\Gamma(\alpha)}$$
- Integrated hazard $$H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(\beta^{\kappa}[y-\mu]^{\kappa},\alpha)]$$
- Hazard $$h(y) = \frac{\kappa(y-\mu)^{\kappa\alpha-1}(\beta^{\kappa})^{\alpha}\mathrm{e}^{-\beta^{\kappa}(y-\mu)^{\kappa}}}{\Gamma(\alpha)-\Gamma(\beta^{\kappa}[y-\mu]^{\kappa},\alpha)}$$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is $\mu < y$, the scale is $\sigma > 0$, the shape is $\alpha > 0$, and the family is $\kappa > 0$. It includes the Weibull distribution ($\alpha = 1$), the half-normal distribution ($\alpha = \frac{1}{2}$, $\kappa = 2$, $\mu = 0$), and the three parameter gamma distribution ($\kappa = 1$).
(Johnson and Kotz, 1970a, p. 197)

Three parameter distributions

- Density $$f(y) = \frac{\kappa y^{\kappa\alpha-1}\mathrm{e}^{-\left(\frac{y}{\sigma}\right)^{\kappa}}}{\sigma^{\kappa\alpha}\Gamma(\alpha)} = \frac{\kappa y^{\kappa\alpha-1}(\beta^{\kappa})^{\alpha}\mathrm{e}^{-\beta^{\kappa}y^{\kappa}}}{\Gamma(\alpha)}$$
- CDF $$F(y) = \frac{\Gamma(\beta^{\kappa}y^{\kappa},\alpha)}{\Gamma(\alpha)}$$
- Survival function $$S(y) = \frac{\Gamma(\alpha)-\Gamma(\beta^{\kappa}y^{\kappa},\alpha)}{\Gamma(\alpha)}$$
- Integrated hazard $$H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(\beta^{\kappa}y^{\kappa},\alpha)]$$
- Hazard $$h(y) = \frac{\kappa y^{\kappa\alpha-1}(\beta^{\kappa})^{\alpha}\mathrm{e}^{-\beta^{\kappa}y^{\kappa}}}{\Gamma(\alpha)-\Gamma(\beta^{\kappa}y^{\kappa},\alpha)}$$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, the shape is $\alpha > 0$, and the family is $\kappa > 0$.
(Johnson and Kotz, 1970a, p. 197)

- Density $f(y) \;=\; \frac{\kappa(y-\mu)^{\kappa\alpha-1}\mathrm{e}^{-(y-\mu)^{\kappa}}}{\Gamma(\alpha)}$
- CDF $F(y) \;=\; \frac{\Gamma([y-\mu]^{\kappa},\alpha)}{\Gamma(\alpha)}$
- Survival function $S(y) \;=\; \frac{\Gamma(\alpha)-\Gamma([y-\mu]^{\kappa},\alpha)}{\Gamma(\alpha)}$
- Integrated hazard $H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma([y-\mu]^{\kappa},\alpha)]$
- Hazard $h(y) \;=\; \frac{\kappa(y-\mu)^{\kappa\alpha-1}\mathrm{e}^{-(y-\mu)^{\kappa}}}{\Gamma(\alpha)-\Gamma([y-\mu]^{\kappa},\alpha)}$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is $\mu < y$, the scale is set to unity ($\sigma = 1$), the shape is $\alpha > 0$, and the family is $\kappa > 0$.
(Johnson and Kotz, 1970a, p. 197)

Two parameter distribution

- Density $f(y) \;=\; \frac{\kappa y^{\kappa\alpha-1}\mathrm{e}^{-y^{\kappa}}}{\Gamma(\alpha)}$
- CDF $F(y) \;=\; \frac{\Gamma(y^{\kappa},\alpha)}{\Gamma(\alpha)}$
- Survival function $S(y) \;=\; \frac{\Gamma(\alpha)-\Gamma(y^{\kappa},\alpha)}{\Gamma(\alpha)}$
- Integrated hazard $H(y) = \ln[\Gamma(\alpha)] - \ln[\Gamma(\alpha) - \Gamma(y^{\kappa},\alpha)]$
- Hazard $h(y) \;=\; \frac{\kappa y^{\kappa\alpha-1}\mathrm{e}^{-y^{\kappa}}}{\Gamma(\alpha)-\Gamma(y^{\kappa},\alpha)}$

The responses y are positive and $\beta = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), the shape is $\alpha > 0$, and the family is $\kappa > 0$.
(Johnson and Kotz, 1970a, p. 197)

INVERSE GAUSSIAN DISTRIBUTION

Two parameter distribution

- Density
$$f(y) = \frac{e^{-\frac{(y-\mu)^2}{2\sigma\mu^2 y}}}{\sqrt{2\pi\sigma y^3}} = \sqrt{\frac{\psi}{2\pi y^3}} e^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi y+\frac{\psi}{y}\right)} \qquad \text{(B.9)}$$
- CDF
$$F(y) = 1 - \frac{1-\Phi\left(\frac{y-\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\psi\phi}}\Phi\left(-\frac{y+\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)}{\sqrt{\frac{\psi}{2\pi y^3}}}$$
- Survival function
$$S(y) = \frac{1-\Phi\left(\frac{y-\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\psi\phi}}\Phi\left(-\frac{y+\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)}{\sqrt{\frac{\psi}{2\pi y^3}}}$$
- Integrated hazard
$$H(y) = \frac{1}{2}\ln[\psi] - \frac{1}{2}\ln[2\pi y^3] - \ln\left[1-\Phi\left(\frac{y-\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\psi\phi}}\Phi\left(-\frac{y+\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)\right]$$
- Hazard
$$h(y) = \frac{\psi}{2\pi y^3}\frac{e^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi y+\frac{\psi}{y}\right)}}{1-\Phi\left(\frac{y-\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\psi\phi}}\Phi\left(-\frac{y+\sqrt{\frac{\psi}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)}$$

The responses y are positive, $\phi = \frac{1}{\sigma\mu^2}$, and $\psi = \frac{1}{\sigma}$. The location is $\mu > 0$ and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970a, pp. 138–139)

One parameter distributions

- Density
$$f(y) = \frac{e^{-\frac{(y-1)^2}{2\sigma y}}}{\sqrt{2\pi\sigma y^3}} = \sqrt{\frac{\phi}{2\pi y^3}} e^{\phi-\frac{\phi}{2}\left(y+\frac{1}{y}\right)}$$
- CDF
$$F(y) = 1 - \sqrt{\frac{2\pi y^3}{\phi}}\left[1-\Phi(y-1)-e^{2\phi}\Phi(-y-1)\right]$$
- Survival function
$$S(y) = \sqrt{\frac{2\pi y^3}{\phi}}\left[1-\Phi(y-1)-e^{2\phi}\Phi(-y-1)\right]$$
- Integrated hazard
$$H(y) = \frac{1}{2}\ln[\phi] - \frac{1}{2}\ln[2\pi y^3] - \ln\left[1-\Phi(y-1)-e^{2\phi}\Phi(-y-1)\right]$$
- Hazard
$$h(y) = \frac{\phi e^{\phi-\frac{\phi}{2}\left(y+\frac{1}{y}\right)}}{2\pi y^3\left[1-\Phi(y-1)-e^{2\phi}\Phi(-y-1)\right]}$$

The responses y are positive and $\phi = \frac{1}{\sigma}$. The location is set to unity ($\mu = 1$) and the scale is $\sigma > 0$. This is the standard from of the distribution which is called the Wald distribution.

(Johnson and Kotz, 1970a, p. 138)

- Density $$f(y) = \frac{e^{-\frac{(y-\mu)^2}{2\mu^2 y}}}{\sqrt{2\pi y^3}} = \frac{e^{\sqrt{\phi}-\frac{1}{2}\left(\phi y+\frac{1}{y}\right)}}{\sqrt{2\pi y^3}}$$
- CDF
$$F(y) = 1-\sqrt{2\pi y^3}\left[1-\Phi\left(\frac{y-\frac{1}{\sqrt{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\phi}}\Phi\left(-\frac{y+\frac{1}{\sqrt{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)\right]$$
- Survival function
$$S(y) = \sqrt{2\pi y^3}\left[1-\Phi\left(\frac{y-\frac{1}{\sqrt{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\phi}}\Phi\left(-\frac{y+\frac{1}{\sqrt{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)\right]$$
- Integrated hazard
$$H(y) = -\frac{1}{2}\ln[2\pi y^3]-\ln\left[1-\Phi\left(\frac{y-\sqrt{\frac{1}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{\phi}}\Phi\left(-\frac{y+\sqrt{\frac{1}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)\right]$$
- Hazard $$h(y) = \frac{e^{\sqrt{\phi}-\frac{1}{2}\left(\phi y+\frac{1}{y}\right)}}{2\pi y^3\left[1-\Phi\left(\frac{y-\sqrt{\frac{1}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)-e^{2\sqrt{1\phi}}\Phi\left(-\frac{y+\sqrt{\frac{1}{\phi}}}{\sqrt{\frac{y}{\phi}}}\right)\right]}$$

The responses y are positive and $\phi = \frac{1}{\mu^2}$. The location is $\mu > 0$ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970a, pp. 138–139)

GENERALIZED INVERSE GAUSSIAN DISTRIBUTION

Three parameter distribution

- Density $$f(y) = \frac{y^{\kappa-1}e^{-\frac{1}{2\sigma}\left(\frac{1}{y}+\frac{y}{\mu^2}\right)}}{2\mu^\kappa K\left(\frac{1}{\sigma\mu},|\kappa|\right)}$$
$$= \frac{y^{\kappa-1}e^{-\frac{y\phi}{2}-\frac{\psi}{2y}}}{2\left(\frac{\psi}{\phi}\right)^{\frac{\kappa}{2}}K\left(\sqrt{\phi\psi},|\kappa|\right)} \tag{B.10}$$

The responses y are positive, $\phi = \frac{1}{\sigma\mu^2}$, and $\psi = \frac{1}{\sigma}$. The location is $\mu > 0$, the scale is $\sigma > 0$, and the family is κ. It includes the inverse Gaussian distribution ($\kappa = q-\frac{1}{2}$),the random walk distribution or reciprocal inverse Gaussian distribution ($\kappa = \frac{1}{2}$), the two parameter gamma distribution ($\sigma = \infty$, $\kappa > 0$), the two parameter reciprocal gamma distribution ($\sigma = \infty$, $\kappa < 0$), and the hyperbola distribution ($\kappa = 0$).
(Jørgensen, 1982, p. 1)

Two parameter distributions

- Density $$f(y) = \frac{y^{\kappa-1}\mathrm{e}^{-\frac{1}{2\sigma}\left(\frac{1}{y}+y\right)}}{2K\left(\frac{1}{\sigma},|\kappa|\right)}$$ $$= \frac{y^{\kappa-1}\mathrm{e}^{-\frac{y\phi}{2}-\frac{\phi}{2y}}}{2K(\phi,|\kappa|)}$$

The responses y are positive and $\phi = \frac{1}{\sigma}$. The location is set to unity ($\mu = 1$), scale is $\sigma > 0$, and the family is κ.
(Jørgensen, 1982, p. 1)

- Density $$f(y) = \frac{y^{\kappa-1}\mathrm{e}^{-\frac{1}{2}\left(\frac{1}{y}+\frac{y}{\mu^2}\right)}}{2\mu^\kappa K\left(\frac{1}{\mu},|\kappa|\right)}$$ $$= \frac{\phi^{\frac{\kappa}{2}}y^{\kappa-1}\mathrm{e}^{-\frac{y\phi}{2}-\frac{1}{2y}}}{2K\left(\sqrt{\phi},|\kappa|\right)}$$

The responses y are positive and $\phi = \frac{1}{\mu^2}$. The location is $\mu > 0$, the scale is set to unity ($\sigma = 1$), and the family is κ.
(Jørgensen, 1982, p. 1)

RANDOM WALK DISTRIBUTION

Two parameter distribution

- Density $$f(y) = \frac{\mathrm{e}^{-\frac{(y\mu-1)^2}{2\sigma\mu^2 y}}}{\sqrt{2\pi\sigma y}}$$ $$= \sqrt{\frac{\phi}{2\pi y}}\mathrm{e}^{\sqrt{\psi\phi}-\frac{1}{2}\left(\phi y+\frac{\psi}{y}\right)}$$

The responses y are positive, $\phi = \frac{1}{\sigma}$, and $\psi = \frac{1}{\sigma\mu^2}$. The location is $\mu > 0$ and the scale is $\sigma > 0$. It is also sometimes called a reciprocal inverse Gaussian distribution.
(Johnson and Kotz, 1970a, pp. 140 [corrected] and 149)

One parameter distributions

- Density $$f(y) = \frac{\mathrm{e}^{-\frac{(y-1)^2}{2\sigma y}}}{\sqrt{2\pi\sigma y}}$$ $$= \sqrt{\frac{\phi}{2\pi y}}\mathrm{e}^{\phi-\frac{\phi}{2}\left(y+\frac{1}{y}\right)} \tag{B.11}$$

The responses y are positive and $\phi = \frac{1}{\sigma}$. The location is set to unity ($\mu = 1$) and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970a, pp. 140 [corrected] and 149)

- Density $$f(y) = \frac{\mathrm{e}^{-\frac{(y\mu-1)^2}{2\mu^2 y}}}{\sqrt{2\pi y}}$$ $$= \sqrt{\frac{1}{2\pi y}}\mathrm{e}^{\sqrt{\psi}-\frac{1}{2}\left(y+\frac{\psi}{y}\right)}$$

The responses y are positive and $\psi = \frac{1}{\sigma}$. The location is $\mu > 0$ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970a, pp. 140 [corrected] and 149)

WEIBULL DISTRIBUTION

Three parameter distribution

- Density $$f(y) = \alpha\sigma^{-\alpha}(y-\mu)^{\alpha-1}\mathrm{e}^{-\left(\frac{y-\mu}{\sigma}\right)^{\alpha}} = \alpha\theta^{\alpha}(y-\mu)^{\alpha-1}\mathrm{e}^{-[\theta(y-\mu)]^{\alpha}}$$
- CDF $$F(y) = 1 - e^{-[\theta(y-\mu)]^{\alpha}}$$
- Survival function $$S(y) = e^{-[\theta(y-\mu)]^{\alpha}}$$
- Integrated hazard $$H(y) = [\theta(y-\mu)]^{\alpha}$$
- Hazard $$h(y) = \alpha\theta^{\alpha}(y-\mu)^{\alpha-1}$$

The responses y are positive and $\theta = \frac{1}{\sigma}$. The location is $\mu < y$, the scale is $\sigma > 0$, and the shape is $\alpha > 0$.
(Johnson and Kotz, 1970a, p. 250)

Two parameter distributions

- Density $$f(y) = \alpha\sigma^{-\alpha}y^{\alpha-1}\mathrm{e}^{-\left(\frac{y}{\sigma}\right)^{\alpha}} = \alpha\theta^{\alpha}y^{\alpha-1}\mathrm{e}^{-(\theta y)^{\alpha}} \qquad \text{(B.12)}$$
- CDF $$F(y) = 1 - e^{-(\theta y)^{\alpha}}$$
- Survival function $$S(y) = e^{-(\theta y)^{\alpha}} \qquad \text{(B.13)}$$
- Integrated hazard $$H(y) = (\theta y)^{\alpha}$$
- Hazard $$h(y) = \alpha\theta^{\alpha}y^{\alpha-1}$$

The responses y are positive and $\theta = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$), scale is $\sigma > 0$, and the shape is $\alpha > 0$.
(Johnson and Kotz, 1970a, p. 250)

- Density $$f(y) = \alpha(y-\mu)^{\alpha-1}\mathrm{e}^{-(y-\mu)^{\alpha}}$$
- CDF $$F(y) = 1 - e^{-(y-\mu)^{\alpha}}$$
- Survival function $$S(y) = e^{-(y-\mu)^{\alpha}}$$
- Integrated hazard $$H(y) = (y-\mu)^{\alpha}$$
- Hazard $$h(y) = \alpha(y-\mu)^{\alpha-1}$$

The responses y are positive. The location is $\mu < y$, the scale is set to unity ($\sigma = 1$), and the shape is $\alpha > 0$.
(Johnson and Kotz, 1970a, p. 250)

One parameter distribution

- Density $f(y) \;=\; \alpha y^{\alpha-1}\mathrm{e}^{-y^{\alpha}}$
- CDF $F(y) \;=\; 1 - e^{-y^{\alpha}}$
- Survival function $S(y) \;=\; e^{-y^{\alpha}}$
- Integrated hazard $H(y) = \; y^{\alpha}$ (B.14)
- Hazard $h(y) \;=\; \alpha y^{\alpha-1}$

The responses y are positive. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the shape is $\alpha > 0$. This is the standard form of the distribution. (Johnson and Kotz, 1970a, p. 250)

GUMBEL DISTRIBUTION

Two parameter distribution

- Density $f(y) \;=\; \dfrac{\mathrm{e}^{-\frac{y-\mu}{\sigma}-\mathrm{e}^{-\frac{y-\mu}{\sigma}}}}{\sigma}$
  $\;=\; \beta\mathrm{e}^{-\beta(y-\mu)-\mathrm{e}^{-\beta(y-\mu)}}$ (B.15)
- CDF $F(y) \;=\; \mathrm{e}^{-\mathrm{e}^{-\beta(y-\mu)}}$ (B.16)
- Survival function $S(y) \;=\; 1 - \mathrm{e}^{-\mathrm{e}^{-\beta(y-\mu)}}$
- Integrated hazard $H(y) = -\ln\left[1 - \mathrm{e}^{-\mathrm{e}^{-\beta(y-\mu)}}\right]$
- Hazard $h(y) \;=\; \dfrac{\beta\mathrm{e}^{-\beta(y-\mu)-\mathrm{e}^{-\beta(y-\mu)}}}{1-\mathrm{e}^{-\mathrm{e}^{-\beta(y-\mu)}}}$

The location is μ, the scale is $\sigma > 0$, and $\beta = \frac{1}{\sigma}$. This distribution belongs to the extreme value family.
(Johnson and Kotz, 1970a, pp. 272 and 277; Agresti, 1990, p. 105)

One parameter distributions

- Density $f(y) \;=\; \dfrac{\mathrm{e}^{-\frac{y}{\sigma}-\mathrm{e}^{-\frac{y}{\sigma}}}}{\sigma}$
  $\;=\; \beta\mathrm{e}^{-\beta y-\mathrm{e}^{-\beta y}}$
- CDF $F(y) \;=\; \mathrm{e}^{-\mathrm{e}^{-\beta y}}$
- Survival function $S(y) \;=\; 1 - \mathrm{e}^{-\mathrm{e}^{-\beta y}}$
- Integrated hazard $H(y) = -\ln\left[1 - \mathrm{e}^{-\mathrm{e}^{-\beta y}}\right]$
- Hazard $h(y) \;=\; \dfrac{\beta\mathrm{e}^{-\beta y-\mathrm{e}^{-\beta y}}}{1-\mathrm{e}^{-\mathrm{e}^{-\beta y}}}$

The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$ and $\beta = \frac{1}{\sigma}$.
(Johnson and Kotz, 1970a, pp. 272 and 277; Agresti, 1990, p. 105)

- Density $f(y) = \mathrm{e}^{-(y-\mu)-\mathrm{e}^{-(y-\mu)}}$
- CDF $F(y) = \mathrm{e}^{-\mathrm{e}^{-(y-\mu)}}$
- Survival function $S(y) = 1 - \mathrm{e}^{-\mathrm{e}^{-(y-\mu)}}$
- Integrated hazard $H(y) = -\ln\left[1 - \mathrm{e}^{-\mathrm{e}^{-(y-\mu)}}\right]$
- Hazard $h(y) = \frac{\mathrm{e}^{-(y-\mu)-\mathrm{e}^{-(y-\mu)}}}{1-\mathrm{e}^{-\mathrm{e}^{-(y-\mu)}}}$

The location is μ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970a, pp. 272 and 277; Agresti, 1990, p. 105)

Zero parameter distribution

- Density $f(y) = \mathrm{e}^{-y-\mathrm{e}^{-y}}$
- CDF $F(y) = \mathrm{e}^{-\mathrm{e}^{-y}}$ (B.17)
- Survival function $S(y) = 1 - \mathrm{e}^{-\mathrm{e}^{-y}}$
- Integrated hazard $H(y) = -\ln\left[1 - \mathrm{e}^{-\mathrm{e}^{-y}}\right]$
- Hazard $h(y) = \frac{\mathrm{e}^{-y-\mathrm{e}^{-y}}}{1-\mathrm{e}^{-\mathrm{e}^{-y}}}$

The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$). This is the standard form of the distribution.
(Johnson and Kotz, 1970a, p. 277; Agresti, 1990, p. 105)

GOMPERTZ DISTRIBUTION

Two parameter distribution

- Density $f(y) = \frac{\mathrm{e}^{\frac{y-\mu}{\sigma}-\mathrm{e}^{\frac{y-\mu}{\sigma}}}}{\sigma}$

  $= \epsilon\theta^{\epsilon}\mathrm{e}^{\epsilon y-(\theta \mathrm{e}^{y})^{\epsilon}}$
- CDF $F(y) = 1 - \mathrm{e}^{-(\theta \mathrm{e}^{y})^{\epsilon}}$
- Survival function $S(y) = \mathrm{e}^{-(\theta \mathrm{e}^{y})^{\epsilon}}$ (B.18)
- Integrated hazard $H(y) = (\theta \mathrm{e}^{y})^{\epsilon}$
- Hazard $h(y) = \epsilon\theta^{\epsilon}\mathrm{e}^{\epsilon y}$

The responses y are positive, $\epsilon = \frac{1}{\sigma}$, and $\theta = \mathrm{e}^{-\mu}$. The location is $\mu > 0$ and the scale is $\sigma > 0$. This distribution belongs to the extreme value family.
(Lindsey, 1999, p. 28)

One parameter distributions

- Density $f(y) = \frac{e^{\frac{y}{\sigma}-e^{\frac{y}{\sigma}}}}{\sigma}$
  $= \epsilon e^{\epsilon y - e^{\epsilon y}}$
- CDF $F(y) = 1 - e^{-e^{\epsilon y}}$
- Survival function $S(y) = e^{-e^{\epsilon y}}$
- Integrated hazard $H(y) = e^{\epsilon y}$
- Hazard $h(y) = \epsilon e^{\epsilon y}$

The responses y are positive and $\epsilon = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$) and the scale is $\sigma > 0$.
(Lindsey, 1999, p. 28)

- Density $f(y) = e^{(y-\mu)-e^{y-\mu}}$
  $= \theta e^{y-\theta e^{y}}$
- CDF $F(y) = 1 - e^{-\theta e^{y}}$
- Survival function $S(y) = e^{-\theta e^{y}}$
- Integrated hazard $H(y) = \theta e^{y}$
- Hazard $h(y) = \theta e^{y}$

The responses y are positive and $\theta = e^{-\mu}$. The location is $\mu > 0$ and the scale is set to unity ($\sigma = 1$). Note that both the hazard and the integrated hazard have the same form.
(Lindsey, 1999, p. 28)
Zero parameter distribution

- Density $f(y) = e^{y-e^{y}}$
- CDF $F(y) = 1 - e^{-e^{y}}$ (B.19)
- Survival function $S(y) = e^{-e^{y}}$ (B.20)
- Integrated hazard $H(y) = e^{y}$ (B.21)
- Hazard $h(y) = e^{y}$

The responses y are positive. The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$). Note that both the hazard and the integrated hazard have the same form. This is the standard form of the distribution.
(Lindsey, 1999, p. 28)

GENERALIZED GOMPERTZ DISTRIBUTION

Three parameter distribution

$$
\begin{aligned}
&\bullet\ \text{Density} && f(y) = \frac{\kappa y^{\kappa-1} e^{\frac{y^\kappa-\mu}{\sigma} - e^{\frac{y^\kappa-\mu}{\sigma}}}}{\sigma} \\
& && = \epsilon\kappa y^{\kappa-1}\theta^\epsilon e^{\epsilon y^\kappa - \theta^\epsilon e^{\epsilon y^\kappa}} \\
&\bullet\ \text{CDF} && F(y) = 1 - e^{-\theta^\epsilon e^{\epsilon y^\kappa}} \\
&\bullet\ \text{Survival function} && S(y) = e^{-\theta^\epsilon e^{\epsilon y^\kappa}} \\
&\bullet\ \text{Integrated hazard} && H(y) = \theta^\epsilon e^{\epsilon y^\kappa} \\
&\bullet\ \text{Hazard} && h(y) = \epsilon\kappa y^{\kappa-1}\theta^\epsilon e^{\epsilon y^\kappa}
\end{aligned}
\tag{B.22}
$$

The responses y are positive, $\epsilon = \frac{1}{\sigma}$, and $\theta = e^{-\mu}$. The location is $\mu > 0$, the scale is $\sigma > 0$, and the family is $\kappa > 0$. This distribution belongs to the generalized extreme value family. It includes the two parameter Gompertz distribution ($\kappa = 1$).
(Lindsey, 1999, p. 293 [corrected])

Two parameter distributions

$$
\begin{aligned}
&\bullet\ \text{Density} && f(y) = \frac{\kappa y^{\kappa-1} e^{\frac{y^\kappa}{\sigma} - e^{\frac{y^\kappa}{\sigma}}}}{\sigma} \\
& && = \epsilon\kappa y^{\kappa-1} e^{\epsilon y^\kappa - e^{\epsilon y^\kappa}} \\
&\bullet\ \text{CDF} && F(y) = 1 - e^{-e^{\epsilon y^\kappa}} \\
&\bullet\ \text{Survival function} && S(y) = e^{-e^{\epsilon y^\kappa}} \\
&\bullet\ \text{Integrated hazard} && H(y) = e^{\epsilon y^\kappa} \\
&\bullet\ \text{Hazard} && h(y) = \epsilon\kappa y^{\kappa-1} e^{\epsilon y^\kappa}
\end{aligned}
$$

The responses y are positive and $\epsilon = \frac{1}{\sigma}$. The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, and the family is $\kappa > 0$.
(Lindsey, 1999, p. 293 [corrected])

$$
\begin{aligned}
&\bullet\ \text{Density} && f(y) = \kappa y^{\kappa-1} e^{(y^\kappa - \mu) - e^{y^\kappa - \mu}} \\
& && = \kappa y^{\kappa-1}\theta e^{y^\kappa - \theta e^{y^\kappa}} \\
&\bullet\ \text{CDF} && F(y) = 1 - e^{-\theta e^{y^\kappa}} \\
&\bullet\ \text{Survival function} && S(y) = e^{-\theta e^{y^\kappa}} \\
&\bullet\ \text{Integrated hazard} && H(y) = \theta e^{y^\kappa} \\
&\bullet\ \text{Hazard} && h(y) = \kappa y^{\kappa-1}\theta e^{y^\kappa}
\end{aligned}
$$

The responses y are positive and $\theta = e^{-\mu}$. The location is $\mu > 0$, the scale is set to unity ($\sigma = 1$), and the family is $\kappa > 0$.
(Lindsey, 1999, p. 293 [corrected])

PARETO DISTRIBUTION

Three parameter distribution

- Density $f(y) = \frac{\alpha\gamma}{(\beta+y)^{\alpha+1}}$
- CDF $F(y) = 1 - \frac{\gamma}{(\beta+y)^{\alpha}}$
- Survival function $S(y) = \frac{\gamma}{(\beta+y)^{\alpha}}$
- Integrated hazard $H(y) = \alpha \ln[\beta+y] - \ln[\gamma]$
- Hazard $h(y) = \frac{\alpha}{\beta+y}$

The responses y are positive and greater than $\gamma^{\frac{1}{\alpha}} - \beta$. The parameters α, β, and γ are positive. This form is also called the Lomax distribution.
(Johnson and Kotz, 1970a, p. 234)

Two parameter distributions

- Density $f(y) = \frac{\alpha\beta^{\alpha}}{(\beta+y)^{\alpha+1}}$ (B.23)
- CDF $F(y) = 1 - \left(\frac{\beta}{\beta+y}\right)^{\alpha}$
- Survival function $S(y) = \left(\frac{\beta}{\beta+y}\right)^{\alpha}$ (B.24)
- Integrated hazard $H(y) = \alpha \ln[\beta+y] - \alpha \ln[\beta]$
- Hazard $h(y) = \frac{\alpha}{\beta+y}$

The responses y are positive as well as the parameters α and β, and $\gamma = \beta^{\alpha}$. This form is obtained from the mixture of a one parameter exponential with a two parameter gamma mixing distribution.
(Johnson and Kotz, 1970a, pp. 233–234; Cox and Oakes, 1984, pp. 19–20)

- Density $f(y) = \frac{\alpha}{(\beta+y)^{\alpha+1}}$
- CDF $F(y) = 1 - (\beta+y)^{-\alpha}$
- Survival function $S(y) = (\beta+y)^{-\alpha}$
- Integrated hazard $H(y) = \alpha \ln[\beta+y]$
- Hazard $h(y) = \frac{\alpha}{\beta+y}$

The responses y are positive and greater than $1-\beta$, the parameters α and β are positive, and the parameter γ is set to unity.
(Johnson and Kotz, 1970a, p. 234)

- Density $f(y) = \frac{\gamma}{(\beta+y)^2}$
- CDF $F(y) = 1 - \frac{\gamma}{\beta+y}$
- Survival function $S(y) = \frac{\gamma}{\beta+y}$
- Integrated hazard $H(y) = \ln[\beta+y] - \ln[\gamma]$
- Hazard $h(y) = \frac{1}{\beta+y}$

The responses y are positive and greater than $\gamma-\beta$, the parameters β and γ are positive, and the parameter α is set to unity.
(Johnson and Kotz, 1970a, p. 234)

- Density $f(y) = \frac{\alpha\gamma}{y^{\alpha+1}}$
  $= \frac{\alpha\kappa^\alpha}{y^{\alpha+1}}$
- CDF $F(y) = 1 - \left(\frac{\kappa}{y}\right)^\alpha$ (B.25)
- Survival function $S(y) = \left(\frac{\kappa}{y}\right)^\alpha$
- Integrated hazard $H(y) = \alpha\ln[y] - \alpha\ln[\kappa]$
- Hazard $h(y) = \frac{\alpha}{y}$

The responses y are positive and greater than the parameter $\gamma^{\frac{1}{\alpha}}$ $(0 < \kappa \leq y)$. The parameters α and γ must be positive, the parameter β is set to zero, and $\gamma = \kappa^\alpha$. This is the original form of the distribution.
(Johnson and Kotz, 1970a, p. 234)

One parameter distributions

- Density $f(y) = \frac{\beta}{(\beta+y)^2}$
- CDF $F(y) = 1 - \frac{\beta}{\beta+y}$
- Survival function $S(y) = \frac{\beta}{\beta+y}$
- Integrated hazard $H(y) = \ln[\beta+y] - \ln[\beta]$
- Hazard $h(y) = \frac{1}{\beta+y}$

The responses y are positive as well as the parameter β, the parameter α is set to unity, and $\gamma = \beta^\alpha$.
(Johnson and Kotz, 1970a, pp. 233–234; Cox and Oakes, 1984, pp. 19–20)

- Density $f(y) = \frac{\alpha}{(1+y)^{\alpha+1}}$
- CDF $F(y) = 1-(1+y)^{-\alpha}$
- Survival function $S(y) = (1+y)^{-\alpha}$
- Integrated hazard $H(y) = \alpha \ln[1+y]$
- Hazard $h(y) = \frac{\alpha}{1+y}$

The responses y are positive as well as the parameter α, and the parameters β and γ are set to unity.
(Johnson and Kotz, 1970a, pp. 233–234; Cox and Oakes, 1984, pp. 19–20)

- Density $f(y) = (\beta+y)^{-2}$
- CDF $F(y) = 1-\frac{1}{\beta+y}$
- Survival function $S(y) = \frac{1}{\beta+y}$
- Integrated hazard $H(y) = \ln[\beta+y]$
- Hazard $h(y) = \frac{1}{\beta+y}$

The responses y are positive and greater than $1-\beta$, the parameters β is positive, and the parameters α and γ are set to unity.
(Johnson and Kotz, 1970a, p. 234)

- Density $f(y) = \frac{\gamma}{y^2}$
- CDF $F(y) = 1-\frac{\gamma}{y}$
- Survival function $S(y) = \frac{\gamma}{y}$
- Integrated hazard $H(y) = \ln[y]-\ln[\gamma]$
- Hazard $h(y) = \frac{1}{y}$

The responses y are positive and greater than the parameter γ ($0<\gamma\leq y$), the parameter α is set to unity, and the parameter β is set to zero.
(Johnson and Kotz, 1970a, p. 234)

LOGISTIC DISTRIBUTION

Two parameter distribution

- Density $$f(y) = \frac{e^{-\frac{y-\mu}{\sigma}}}{\sigma\left(1+e^{-\frac{y-\mu}{\sigma}}\right)^2} \tag{B.26}$$
- CDF $$F(y) = \frac{1}{1+e^{-\frac{y-\mu}{\sigma}}} \tag{B.27}$$
- Survival function $$S(y) = \frac{1}{1+e^{\frac{y-\mu}{\sigma}}}$$
- Integrated hazard $$H(y) = \ln\left[1+e^{\frac{y-\mu}{\sigma}}\right]$$
- Hazard $$h(y) = \frac{1}{\sigma\left(1+e^{-\frac{y-\mu}{\sigma}}\right)}$$

The location is μ and the scale is $\sigma > 0$. This form is obtained from the mixture of a two parameter Gumbel distribution with a one parameter exponential mixing distribution.
(Johnson and Kotz, 1970b, pp. 1 and 3)

One parameter distributions

- Density $$f(y) = \frac{e^{-\frac{y}{\sigma}}}{\sigma\left(1+e^{-\frac{y}{\sigma}}\right)^2}$$
- CDF $$F(y) = \frac{1}{1+e^{-\frac{y}{\sigma}}}$$
- Survival function $$S(y) = \frac{1}{1+e^{\frac{y}{\sigma}}} \tag{B.28}$$
- Integrated hazard $$H(y) = \ln\left[1+e^{\frac{y}{\sigma}}\right]$$
- Hazard $$h(y) = \frac{1}{\sigma\left(1+e^{-\frac{y}{\sigma}}\right)}$$

The location is set to zero ($\mu = 0$) and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970b, p. 1)

- Density $$f(y) = \frac{e^{\mu-y}}{(1+e^{\mu-y})^2}$$
- CDF $$F(y) = \frac{1}{1+e^{\mu-y}}$$
- Survival function $$S(y) = \frac{1}{1+e^{y-\mu}}$$
- Integrated hazard $$H(y) = \ln\left[1+e^{y-\mu}\right]$$
- Hazard $$h(y) = \frac{1}{(1+e^{\mu-y})}$$

The location is μ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970b, p. 1)

Zero parameter distribution

- Density $f(y) \ = \ \frac{\mathrm{e}^{-y}}{(1+\mathrm{e}^{-y})^2}$
- CDF $F(y) \ = \ \frac{1}{1+\mathrm{e}^{-y}}$ (B.29)
- Survival function $S(y) \ = \ \frac{1}{1+\mathrm{e}^{y}}$
- Integrated hazard $H(y) = \ \ln[1+\mathrm{e}^{y}]$
- Hazard $h(y) \ = \ \frac{1}{(1+\mathrm{e}^{-y})}$

The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$). This is the standard form of the distribution.
(Johnson and Kotz, 1970b, p. 3)

GENERALIZED LOGISTIC DISTRIBUTION

Four parameter distribution

- Density $f(y) \ = \ \frac{\alpha\beta^{\alpha}\mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\left(\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)^{\alpha+1}}$ (B.30)
- CDF $F(y) \ = \ \left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha}$ (B.31)
- Survival function $S(y) \ = \ 1-\left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha}$
- Integrated hazard $H(y) = \ -\ln\left[1-\left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha}\right]$
- Hazard $h(y) \ = \ \frac{\alpha\beta^{\alpha}\mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\left(\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)\left[\left(\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)^{\alpha}-\beta^{\alpha}\right]}$

The location is μ, the scale is $\sigma > 0$, and the two shape parameters α and β are positive. It includes the two parameter logistic distribution ($\alpha = 1$ and $\beta = \sigma$). This form is obtained from the mixture of a two parameter Gumbel distribution with a two parameter gamma mixing distribution.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

Three parameter distributions

- Density $f(y) \;=\; \frac{\alpha\beta^{\alpha}\,\mathrm{e}^{-\frac{y}{\sigma}}}{\left(\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}\right)^{\alpha+1}}$
- CDF $F(y) \;=\; \left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}}\right)^{\alpha}$
- Survival function $S(y) \;=\; 1-\left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}}\right)^{\alpha}$
- Integrated hazard $H(y) = -\ln\left[1-\left(\frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}}\right)^{\alpha}\right]$
- Hazard $h(y) \;=\; \frac{\alpha\beta^{\alpha}\,\mathrm{e}^{-\frac{y}{\sigma}}}{\left(\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}\right)\left[\left(\beta+\sigma\mathrm{e}^{-\frac{y}{\sigma}}\right)^{\alpha}-\beta^{\alpha}\right]}$

The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, and the two shape parameters α and β are positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $f(y) \;=\; \frac{\alpha\beta^{\alpha}\,\mathrm{e}^{\mu-y}}{(\beta+\mathrm{e}^{\mu-y})^{\alpha+1}}$
- CDF $F(y) \;=\; \left(\frac{\beta}{\beta+\mathrm{e}^{\mu-y}}\right)^{\alpha}$
- Survival function $S(y) \;=\; 1-\left(\frac{\beta}{\beta+\mathrm{e}^{\mu-y}}\right)^{\alpha}$
- Integrated hazard $H(y) = -\ln\left[1-\left(\frac{\beta}{\beta+\mathrm{e}^{\mu-y}}\right)^{\alpha}\right]$
- Hazard $h(y) \;=\; \frac{\alpha\beta^{\alpha}\,\mathrm{e}^{\mu-y}}{(\beta+\mathrm{e}^{\mu-y})[(\beta+\mathrm{e}^{\mu-y})^{\alpha}-\beta^{\alpha}]}$

The location is μ, the scale is set to unity ($\sigma = 1$), and the two shape parameters α and β are positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $f(y) \;=\; \frac{\beta\mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\left(\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)^{2}}$
- CDF $F(y) \;=\; \frac{\beta}{\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}}$
- Survival function $S(y) \;=\; \frac{\sigma}{\sigma+\beta\mathrm{e}^{\frac{y-\mu}{\sigma}}}$
- Integrated hazard $H(y) = \ln\left[\sigma+\beta\mathrm{e}^{\frac{y-\mu}{\sigma}}\right]-\ln[\sigma]$
- Hazard $h(y) \;=\; \frac{\beta}{\sigma\left(\beta+\sigma\mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)}$

The location is μ, the scale is $\sigma > 0$, the first shape parameter is set to unity ($\alpha = 1$) and the second shape parameter β is positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $\quad f(y) \;=\; \frac{\alpha \mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\left(1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)^{\alpha+1}}$
- CDF $\quad F(y) \;=\; \left(\frac{1}{1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha} \qquad$ (B.32)
- Survival function $\quad S(y) \;=\; 1-\left(\frac{1}{1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha}$
- Integrated hazard $\quad H(y) = -\ln\left[1-\left(\frac{1}{1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}}\right)^{\alpha}\right]$
- Hazard $\quad h(y) \;=\; \frac{\alpha \mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\left(1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)\left[\left(1+\sigma \mathrm{e}^{-\frac{y-\mu}{\sigma}}\right)^{\alpha}-1\right]}$

The location is μ, the scale is $\sigma > 0$, the first shape parameter α is positive and the second shape parameter is set to unity ($\beta = 1$).
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

Two parameter distributions

- Density $\quad f(y) \;=\; \frac{\alpha\beta^{\alpha}\mathrm{e}^{-y}}{(\beta+\mathrm{e}^{-y})^{\alpha+1}}$
- CDF $\quad F(y) \;=\; \left(\frac{\beta}{\beta+\mathrm{e}^{-y}}\right)^{\alpha}$
- Survival function $\quad S(y) \;=\; 1-\left(\frac{\beta}{\beta+\mathrm{e}^{-y}}\right)^{\alpha}$
- Integrated hazard $\quad H(y) = -\ln\left[1-\left(\frac{\beta}{\beta+\mathrm{e}^{-y}}\right)^{\alpha}\right]$
- Hazard $\quad h(y) \;=\; \frac{\alpha\beta^{\alpha}\mathrm{e}^{-y}}{(\beta+\mathrm{e}^{-y})[(\beta+\mathrm{e}^{-y})^{\alpha}-\beta^{\alpha}]}$

The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the two shape parameters α and β are positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $\quad f(y) \;=\; \frac{\beta \mathrm{e}^{-\frac{y}{\sigma}}}{\left(\beta+\sigma \mathrm{e}^{-\frac{y}{\sigma}}\right)^{2}}$
- CDF $\quad F(y) \;=\; \frac{\beta}{\beta+\sigma \mathrm{e}^{-\frac{y}{\sigma}}}$
- Survival function $\quad S(y) \;=\; \frac{\sigma}{\sigma+\beta \mathrm{e}^{\frac{y}{\sigma}}}$
- Integrated hazard $\quad H(y) = \ln\left[\sigma+\beta \mathrm{e}^{\frac{y}{\sigma}}\right]-\ln[\sigma]$
- Hazard $\quad h(y) \;=\; \frac{\beta}{\sigma\left(\beta+\sigma \mathrm{e}^{-\frac{y}{\sigma}}\right)}$

The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, the first shape parameter is set to unity ($\alpha = 1$) and the second shape parameter β is positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $f(y) \; = \; \frac{\alpha e^{-\frac{y}{\sigma}}}{\left(1+\sigma e^{-\frac{y}{\sigma}}\right)^{\alpha+1}}$
- CDF $F(y) \; = \; \left(\frac{1}{1+\sigma e^{-\frac{y}{\sigma}}}\right)^{\alpha}$
- Survival function $S(y) \; = \; 1-\left(\frac{1}{1+\sigma e^{-\frac{y}{\sigma}}}\right)^{\alpha}$
- Integrated hazard $H(y) = -\ln\left[1-\left(\frac{1}{1+\sigma e^{-\frac{y}{\sigma}}}\right)^{\alpha}\right]$
- Hazard $h(y) \; = \; \frac{\alpha e^{-\frac{y}{\sigma}}}{\left(1+\sigma e^{-\frac{y}{\sigma}}\right)\left[\left(1+\sigma e^{-\frac{y}{\sigma}}\right)^{\alpha}-1\right]}$

The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, the first shape parameter α is positive and the second shape parameter is set to unity ($\beta = 1$). This form is obtained from the mixture of a two parameter Gompertz distribution with a two parameter gamma mixing distribution.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $f(y) \; = \; \frac{\beta e^{\mu-y}}{(\beta+e^{\mu-y})^{2}}$
- CDF $F(y) \; = \; \frac{\beta}{\beta+e^{\mu-y}}$
- Survival function $S(y) \; = \; \frac{1}{1+\beta e^{y-\mu}}$
- Integrated hazard $H(y) = \ln[1+\beta e^{y-\mu}]$
- Hazard $h(y) \; = \; \frac{\beta}{\beta+e^{\mu-y}}$

The location is μ, the scale is set to unity ($\sigma = 1$), the first shape parameter is set to unity ($\alpha = 1$) and the second shape parameter β is positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

- Density $f(y) \; = \; \frac{\alpha e^{\mu-y}}{(1+e^{\mu-y})^{\alpha+1}}$
- CDF $F(y) \; = \; (1+e^{\mu-y})^{-\alpha}$
- Survival function $S(y) \; = \; 1-(1+e^{\mu-y})^{-\alpha}$
- Integrated hazard $H(y) = -\ln\left[1-(1+e^{\mu-y})^{-\alpha}\right]$
- Hazard $h(y) \; = \; \frac{\alpha e^{\mu-y}}{(1+e^{\mu-y})[(1+e^{\mu-y})^{\alpha}-1]}$

The location is μ, the scale is $\sigma > 0$, the first shape parameter α is positive and the second shape parameter is set equal to the scale ($\beta = \sigma$).
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

One parameter distributions

$$\begin{array}{lrcl} \bullet \text{ Density} & f(y) & = & \frac{\beta \mathrm{e}^{-y}}{(\beta+\mathrm{e}^{-y})^2} \\ \bullet \text{ CDF} & F(y) & = & \frac{\beta}{\beta+\mathrm{e}^{-y}} \\ \bullet \text{ Survival function} & S(y) & = & \frac{1}{1+\beta \mathrm{e}^{y}} \\ \bullet \text{ Integrated hazard} & H(y) & = & \ln[1+\beta \mathrm{e}^{y}] \\ \bullet \text{ Hazard} & h(y) & = & \frac{\beta}{(\beta+\mathrm{e}^{-y})} \end{array} \tag{B.33}$$

The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), the first shape parameter is set to unity ($\alpha = 1$) and the second shape parameter β is positive.
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

$$\begin{array}{lrcl} \bullet \text{ Density} & f(y) & = & \frac{\alpha \mathrm{e}^{-y}}{(1+\mathrm{e}^{-y})^{\alpha+1}} \\ \bullet \text{ CDF} & F(y) & = & (1+\mathrm{e}^{-y})^{-\alpha} \\ \bullet \text{ Survival function} & S(y) & = & 1-(1+\mathrm{e}^{-y})^{-\alpha} \\ \bullet \text{ Integrated hazard} & H(y) & = & -\ln\left[1-(1+\mathrm{e}^{-y})^{-\alpha}\right] \\ \bullet \text{ Hazard} & h(y) & = & \frac{\alpha \mathrm{e}^{-y}}{(1+\mathrm{e}^{-y})[(1+\mathrm{e}^{-y})^{\alpha}-1]} \end{array}$$

The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), the first shape parameter α is positive and the second shape parameter is set to unity ($\beta = 1$).
(Johnson and Kotz, 1970a, p. 289; Johnson and Kotz, 1970b, p. 17 [corrected])

BETA DISTRIBUTION

Four parameter distribution

- Density $f(y) = \frac{1}{B(\alpha,\beta)} \frac{(y-\phi)^{\alpha-1}(\psi-y)^{\beta-1}}{(\psi-\phi)^{\alpha+\beta-1}}$

The responses are included in the interval $\phi \leq y \leq \psi$ and the two shape parameters α and β are positive. It is also sometimes called the power function.
(Johnson and Kotz, 1970b, p. 37)

Two parameter distribution

$$\bullet \text{ Density} \quad f(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)} \tag{B.34}$$

The responses are included in the interval $0 \leq y \leq 1$ as the parameters ϕ and ψ are respectively set to zero and unity. The two shape parameters α and β are positive. This is the standard form of the distribution and is obtained from the mixture of a two parameter gamma distribution with itself.
(Johnson and Kotz, 1970a, p. 195; Johnson and Kotz, 1970b, p. 37)

One parameter distribution

- Density $f(y) = \alpha y^{\alpha-1}$

The responses are included in the interval $0 \leq y \leq 1$ as the parameters ϕ and ψ are respectively set to zero and unity. The first shape parameter is $\alpha > 0$, and the second shape parameter is set to unity ($\beta = 1$). This form is called the standard power function density.
(Johnson and Kotz, 1970b, p. 37)

BURR DISTRIBUTION

Four parameter distribution

- Density $$f(y) = \frac{\theta\alpha\left(\frac{y-\mu}{\sigma}\right)^{\theta-1}}{\sigma\left[1+\left(\frac{y-\mu}{\sigma}\right)^{\theta}\right]^{\alpha+1}} = \frac{\theta\alpha\beta^{\alpha}(y-\mu)^{\theta-1}}{[\beta+(y-\mu)^{\theta}]^{\alpha+1}}$$
- CDF $F(y) = 1 - \left[\frac{\beta}{\beta+(y-\mu)^{\theta}}\right]^{\alpha}$
- Survival function $S(y) = \left[\frac{\beta}{\beta+(y-\mu)^{\theta}}\right]^{\alpha}$
- Integrated hazard $H(y) = \alpha \ln[\beta + (y-\mu)^{\theta}] - \alpha \ln[\beta]$
- Hazard $h(y) = \frac{\theta\alpha(y-\mu)^{\theta-1}}{\beta+(y-\mu)^{\theta}}$

The responses are positive and $\beta = \sigma^{\theta}$. The location is $\mu < y$, the scale is $\sigma > 0$, and the two shape parameters α and θ are both positive. It includes the two parameter Pareto distribution ($\mu = 0$ and $\theta = 1$).
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

Three parameter distributions

- Density $$f(y) = \frac{\theta\alpha\left(\frac{y}{\sigma}\right)^{\theta-1}}{\sigma\left[1+\left(\frac{y}{\sigma}\right)^{\theta}\right]^{\alpha+1}} = \frac{\theta\alpha\beta^{\alpha}y^{\theta-1}}{(\beta+y^{\theta})^{\alpha+1}} \qquad \text{(B.35)}$$
- CDF $$F(y) = 1 - \left[\frac{\beta}{\beta+y^{\theta}}\right]^{\alpha} \qquad \text{(B.36)}$$
- Survival function $$S(y) = \left[\frac{\beta}{\beta+y^{\theta}}\right]^{\alpha} \qquad \text{(B.37)}$$
- Integrated hazard $H(y) = \alpha \ln[\beta + y^{\theta}] - \alpha \ln[\beta]$
- Hazard $h(y) = \frac{\theta\alpha y^{\theta-1}}{\beta+y^{\theta}}$

The responses are positive and $\beta = \sigma^{\theta}$. The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, and the two shape parameters α and θ are both positive. This is

the standard form of the distribution and is obtained from the mixture of a two parameter Weibull with a two parameter gamma mixing distribution.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

- Density $f(y) = \frac{\theta\alpha(y-\mu)^{\theta-1}}{\left[1+(y-\mu)^{\theta}\right]^{\alpha+1}}$
- CDF $F(y) = 1-[1+(y-\mu)^{\theta}]^{-\alpha}$
- Survival function $S(y) = [1+(y-\mu)^{\theta}]^{-\alpha}$
- Integrated hazard $H(y) = \alpha\ln[1+(y-\mu)^{\theta}]$
- Hazard $h(y) = \frac{\theta\alpha(y-\mu)^{\theta-1}}{1+(y-\mu)^{\theta}}$

The responses are positive. The location is $\mu < y$, the scale is set to unity ($\sigma = 1$), and the two shape parameters α and θ are both positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

- Density $f(y) = \frac{\theta\left(\frac{y-\mu}{\sigma}\right)^{\theta-1}}{\sigma\left[1+\left(\frac{y-\mu}{\sigma}\right)^{\theta}\right]^{2}} = \frac{\theta\beta(y-\mu)^{\theta-1}}{[\beta+(y-\mu)^{\theta}]^{2}}$
- CDF $F(y) = \frac{(y-\mu)^{\theta}}{\beta+(y-\mu)^{\theta}}$
- Survival function $S(y) = \frac{\beta}{\beta+(y-\mu)^{\theta}}$
- Integrated hazard $H(y) = \ln[\beta+(y-\mu)^{\theta}]-\ln[\beta]$
- Hazard $h(y) = \frac{\theta(y-\mu)^{\theta-1}}{\beta+(y-\mu)^{\theta}}$

The responses are positive and $\beta = \sigma^{\theta}$. The location is $\mu < y$, the scale is $\sigma > 0$, the first shape parameters α is set to unity and the second shape parameter θ is positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

Two parameter distributions

- Density $f(y) = \frac{\theta\alpha y^{\theta-1}}{\left[1+y^{\theta}\right]^{\alpha+1}}$
- CDF $F(y) = 1-[1+y^{\theta}]^{-\alpha}$
- Survival function $S(y) = [1+y^{\theta}]^{-\alpha}$
- Integrated hazard $H(y) = \alpha\ln[1+y^{\theta}]$
- Hazard $h(y) = \frac{\theta\alpha y^{\theta-1}}{1+y^{\theta}}$

The responses are positive. The location is is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the two shape parameters α and θ are both positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

- Density $f(y) = \frac{\theta\left(\frac{y}{\sigma}\right)^{\theta-1}}{\sigma\left[1+\left(\frac{y}{\sigma}\right)^{\theta}\right]^2}$
  $= \frac{\theta\beta y^{\theta-1}}{(\beta+y^{\theta})^2}$
- CDF $F(y) = \frac{y^{\theta}}{\beta+y^{\theta}}$
- Survival function $S(y) = \frac{\beta}{\beta+y^{\theta}}$
- Integrated hazard $H(y) = \ln[\beta + y^{\theta}] - \ln[\beta]$
- Hazard $h(y) = \frac{\theta y^{\theta-1}}{\beta+y^{\theta}}$

The responses are positive and $\beta = \sigma^{\theta}$. The location is set to zero ($\mu = 0$), the scale is $\sigma > 0$, the first shape parameter α is set to unity and the second shape parameter θ is positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

- Density $f(y) = \frac{\theta(y-\mu)^{\theta-1}}{\left[1+(y-\mu)^{\theta}\right]^2}$
- CDF $F(y) = \frac{(y-\mu)^{\theta}}{1+(y-\mu)^{\theta}}$
- Survival function $S(y) = \frac{1}{1+(y-\mu)^{\theta}}$
- Integrated hazard $H(y) = \ln[1 + (y-\mu)^{\theta}]$
- Hazard $h(y) = \frac{\theta(y-\mu)^{\theta-1}}{1+(y-\mu)^{\theta}}$

The responses are positive. The location is $\mu < y$, the scale is set to unity ($\sigma = 1$), the first shape parameter α is set to unity and the second shape parameter θ is positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

One parameter distribution

- Density $f(y) = \frac{\theta y^{\theta-1}}{\left[1+y^{\theta}\right]^2}$
- CDF $F(y) = \frac{y^{\theta}}{1+y^{\theta}}$ (B.38)
- Survival function $S(y) = \frac{1}{1+y^{\theta}}$
- Integrated hazard $H(y) = \ln[1+y^{\theta}]$
- Hazard $h(y) = \frac{\theta y^{\theta-1}}{1+y^{\theta}}$

The responses are positive. The location is set to unity ($\mu = 1$), the scale is set to unity ($\sigma = 1$), the first shape parameter α is set to unity and the second shape parameter θ is positive.
(Johnson and Kotz, 1970a, pp. 30–31 and 266)

LAPLACE DISTRIBUTION

Two parameter distribution

- Density $f(y) = \frac{e^{-\frac{|y-\mu|}{\sigma}}}{2\sigma}$

$$= \frac{\beta e^{-\beta|y-\mu|}}{2} \qquad \text{(B.39)}$$

The location is μ and the scale is $\sigma > 0$. This form is also called the double exponential.
(Johnson and Kotz, 1970b, pp. 22–23)

One parameter distributions

- Density $f(y) = \frac{e^{-\frac{|y|}{\sigma}}}{2\sigma}$

$$= \frac{\beta e^{-\beta|y|}}{2}$$

The location is set to zero ($\mu = 0$) and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970b, pp. 22–23)

- Density $f(y) = \frac{e^{-|y-\mu|}}{2}$

The location is μ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970b, pp. 22–23)

Zero parameter distribution

- Density $f(y) = \frac{e^{-|y|}}{2}$

The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$). This is the standard form of the distribution.
(Johnson and Kotz, 1970b, pp. 22–23)

SIMPLEX DISTRIBUTION

Two parameter distribution

- Density $f(y) = \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2 y(1-y)\mu^2(1-\mu)^2}}}{\sqrt{2\pi\sigma^2 y^3(1-y)^3}}$ (B.40)

The responses are included in the interval $0 < y < 1$. The location is $0 < \mu < 1$ and the scale is σ.
(Song and Tan, 2000)

One parameter distribution

- Density $f(y) = \frac{e^{-\frac{(y-\mu)^2}{2y(1-y)\mu^2(1-\mu)^2}}}{\sqrt{2\pi y^3(1-y)^3}}$

The responses are included in the interval $0 < y < 1$. The location is $0 < \mu < 1$ and the scale is set to unity ($\sigma = 1$).
(Song and Tan, 2000)

GAUSSIAN DISTRIBUTION

Two parameter distribution

- Density $f(y) = \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$ (B.41)
- CDF $F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$

The location is μ (also theoretical mean) and the scale is $\sigma > 0$ (with theoretical variance σ^2). This is obtained from the mixture of a two parameter Gaussian distribution with itself.
(Johnson and Kotz, 1970a, pp. 40 and 87–88)

One parameter distributions

- Density $f(y) = \frac{e^{-\frac{y^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$
- CDF $F(y) = \Phi\left(\frac{y}{\sigma}\right)$

The location is set to zero ($\mu = 0$) and the scale is $\sigma > 0$.
(Johnson and Kotz, 1970a, p. 40)

- Density $f(y) = \frac{e^{-\frac{(y-\mu)^2}{2}}}{\sqrt{2\pi}}$
- CDF $F(y) = \Phi(y - \mu)$

The location is μ and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1970a, p. 40)

Zero parameter distribution

- Density $f(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$ (B.42)
- CDF $F(y) = \Phi(y)$ (B.43)

The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$). This form is called the standardized Gaussian distribution.
(Johnson and Kotz, 1970a, p. 40)

LOGNORMAL DISTRIBUTION

Three parameter distribution

- Density $f(y) = \frac{e^{-\frac{(\ln[y-\theta]-\mu)^2}{2\sigma^2}}}{(y-\theta)\sqrt{2\pi}\sigma}$ (B.44)

The responses are positive. The location is μ, the scale is σ, and the parameter θ is is positive. This is also called the antilognormal distribution.
(Johnson and Kotz, 1970a, pp. 112–113)

Two parameter distributions

- Density $f(y) = \frac{e^{-\frac{(\ln[y]-\mu)^2}{2\sigma^2}}}{y\sqrt{2\pi}\sigma}$

The responses are positive. The location is μ, the scale is σ, and the parameter θ is set to zero ($\theta = 0$). This is the standard form of the distribution.
(Johnson and Kotz, 1970a, pp. 112–113)

- Density $f(y) = \frac{e^{-\frac{\ln[y-\theta]^2}{2\sigma^2}}}{(y-\theta)\sqrt{2\pi}\sigma}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is σ, and the parameter θ is is positive.
(Johnson and Kotz, 1970a, pp. 112–113)

- Density $f(y) = \frac{e^{-\frac{(\ln[y-\theta]-\mu)^2}{2}}}{(y-\theta)\sqrt{2\pi}}$

The responses are positive. The location is μ, the scale is set to unity ($\sigma = 1$), and the parameter θ is is positive.
(Johnson and Kotz, 1970a, pp. 112–113)

One parameter distributions

- Density $f(y) = \frac{e^{-\frac{\ln[y-\theta]^2}{2}}}{(y-\theta)\sqrt{2\pi}}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the parameter θ is is positive.
(Johnson and Kotz, 1970a, pp. 112–113)

- Density $f(y) = \frac{e^{-\frac{\ln[y]^2}{2\sigma^2}}}{y\sqrt{2\pi}\sigma}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is σ, and the parameter θ is set to zero ($\theta = 0$).
(Johnson and Kotz, 1970a, pp. 112–113)

- Density $f(y) = \frac{e^{-\frac{(\ln[y]-\mu)^2}{2}}}{y\sqrt{2\pi}}$

The responses are positive. The location is μ, the scale is set to unity ($\sigma = 1$), and the parameter θ is set to zero ($\theta = 0$).
(Johnson and Kotz, 1970a, pp. 112–113)

Zero parameter distribution

- Density $\quad f(y) = \frac{e^{-\frac{\ln[y]^2}{2}}}{y\sqrt{2\pi}}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the parameter θ is set to zero ($\theta = 0$).
(Johnson and Kotz, 1970a, pp. 112–113)

B.4 Discrete univariate distributions

POISSON DISTRIBUTION

One parameter distribution

- Probability mass function $\quad \Pr(y) = \frac{\mu^y e^{-\mu}}{y!} \qquad \text{(B.45)}$

The responses y are discrete (non-negative integers, $y = 0, 1, 2\ldots$) and the location is $\mu > 0$. The theoretical mean μ is equal to the theoretical variance.
(Johnson and Kotz, 1969, p. 87)

BINOMIAL DISTRIBUTION

Two parameter distribution

- Density $\quad f(y) = \binom{\alpha}{y}\pi^y(1-\pi)^{\alpha-y} \qquad \text{(B.46)}$

The responses y are discrete (non-negative integers, $y = 0, 1, 2\ldots, \alpha$). The parameter α is also a positive integer and the parameter π must be included in the interval $0 \leq \pi \leq 1$.
(Johnson and Kotz, 1969, p. 50)

One parameter distribution

- Density $\quad f(y) = \pi^y(1-\pi)^{1-y} \qquad \text{(B.47)}$

The responses y is binary ($y = 0, 1$). The parameter α is set to unity ($\alpha = 1$) and the parameter π must be included in the interval $0 \leq \pi \leq 1$. This form is called a Bernoulli distribution.
(Johnson and Kotz, 1969, pp. 50–51)

NEGATIVE BINOMIAL DISTRIBUTION

Two parameter distribution

- Density
$$
\begin{aligned}
f(y) &= \binom{\alpha+y-1}{\alpha-1}\left(\frac{1}{\beta+1}\right)^y\left(\frac{\beta}{\beta+1}\right)^\alpha \\
&= \frac{\beta^\alpha}{B(\alpha,y)y(\beta+1)^{\alpha+y}} \qquad \text{(B.48)}
\end{aligned}
$$

The responses y are discrete (non-negative integers, $y = 0, 1, 2\ldots$). The parameters α and β are both positive. This form is obtained from the mixture of a Poison with a two parameter gamma mixing distribution. It is called sometimes called a binomial waiting-time distribution or a Pólya distribution and a Pascal distribution when α is a positive integer.
(Johnson and Kotz, 1969, pp. 122 and 124–125)

One parameter distribution

- Density $f(y) = \frac{\beta}{(\beta+1)^{y+1}}$

The responses y are discrete (non-negative integers, $y = 0, 1, 2 \ldots$), the parameter α is set to unity ($\alpha = 1$), and the parameter β is positive. This form is called a geometric distribution.
(Johnson and Kotz, 1969, p. 123)

B.5 Multivariate distributions

MULTIVARIATE BURR DISTRIBUTION

Three parameter distribution

- Density

$$f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)}{\Gamma(\alpha)}\left(1+\sum_{j=1}^{m}\beta_j y_j^{\kappa_j}\right)^{-\alpha-m}\prod_{j=1}^{m}\left(\beta_j\kappa_j y_j^{\kappa_j-1}\right) \quad \text{(B.49)}$$

- Survival function $S_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m}\beta_j y_j^{\kappa_j}\right)^{-\alpha}$

The responses are positive. The parameters α, β_j, and κ_j are also all positive. The marginal cdf is a three parameter Burr distribution, $F(y) = 1-(1+\beta y^{\kappa})^{-\alpha}$, corresponding to Equation (B.36).
(Johnson and Kotz, 1972, p. 289; Cook and Johnson, 1981 [method corrected])

One parameter distributions

- Density

$$f_{\mathrm{M}}(\mathbf{y}) = \Gamma(m+1)\left(1+\sum_{j=1}^{m}y_j^{\kappa_j}\right)^{-m-1}\prod_{j=1}^{m}\left(\kappa_j y_j^{\kappa_j-1}\right)$$

- Survival function $S_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m}y_j^{\kappa_j}\right)^{-1}$

The responses are positive. The parameters α and β_j are set to unity. The parameter κ_j is positive. The marginal cdf is a two parameter Burr distribution, $F(y) = \frac{y^{\kappa}}{1+y^{\kappa}}$, corresponding to Equation (B.38).
(Johnson and Kotz, 1972, p. 289; Cook and Johnson, 1981 [method corrected])

MULTIVARIATE LOGISTIC DISTRIBUTION

Two parameter distribution

- Density

$$f_{\mathrm{M}}(\mathbf{y}) = \Gamma(m+1)\left(1+\sum_{j=1}^{m}\mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}\right)^{-m-1}\prod_{j=1}^{m}\frac{\mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}}{\sigma_j} \quad \text{(B.50)}$$

- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m}\mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}\right)^{-1}$

The responses are positive. The location vector contains $\mu_j > 0$ and the scale vector contains $\sigma_j > 0$. The marginal cdf is a two parameter logistic distribution, $F(y) = \frac{1}{1+\mathrm{e}^{-\frac{y-\mu}{\sigma}}}$, corresponding to Equation (B.27).
(Johnson and Kotz, 1972, p. 293 [corrected])

One parameter distributions

- Density $f_{\mathrm{M}}(\mathbf{y}) = \Gamma(m+1)\left(1+\sum_{j=1}^{m} \mathrm{e}^{-\frac{y_j}{\sigma_j}}\right)^{-m-1} \prod_{j=1}^{m} \frac{\mathrm{e}^{-\frac{y_j}{\sigma_j}}}{\sigma_j}$
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m} \mathrm{e}^{-\frac{y_j}{\sigma_j}}\right)^{-1}$

The responses are positive. The location parameters are set to zero ($\mu_j = 0$) and the scale vector contains $\sigma_j > 0$.
(Johnson and Kotz, 1972, p. 293 [corrected])

- Density $f_{\mathrm{M}}(\mathbf{y}) = \Gamma(m+1)\left(1+\sum_{j=1}^{m} \mathrm{e}^{-(y_j-\mu_j)}\right)^{-m-1} \prod_{j=1}^{m} \mathrm{e}^{-(y_j-\mu_j)}$
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m} \mathrm{e}^{-(y_j-\mu_j)}\right)^{-1}$

The responses are positive. The location vector contains $\mu_j > 0$ and the scale parameters are set to unity ($\sigma_j = 1$).
(Johnson and Kotz, 1972, p. 293 [corrected])

Zero parameter distribution

- Density $f_{\mathrm{M}}(\mathbf{y}) = \Gamma(m+1)\left(1+\sum_{j=1}^{m} \mathrm{e}^{-y_j}\right)^{-m-1} \prod_{j=1}^{m} \mathrm{e}^{-y_j}$ (B.51)
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1+\sum_{j=1}^{m} \mathrm{e}^{-y_j}\right)^{-1}$

The responses are positive. The location is set to zero ($\mu = 0$) and the scale is set to unity ($\sigma = 1$).
(Johnson and Kotz, 1972, p. 291; Cook and Johnson, 1981)

MULTIVARIATE GENERALIZED LOGISTIC DISTRIBUTION

Three parameter distributions

- Density
$$f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)\alpha^{\alpha}}{\Gamma(\alpha)}\left(\alpha+\sum_{j=1}^{m} \mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}\right)^{-\alpha-m} \prod_{j=1}^{m} \frac{\mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}}{\sigma_j} \quad \text{(B.52)}$$
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1+\frac{\sum_{j=1}^{m} \mathrm{e}^{-\frac{y_j-\mu_j}{\sigma_j}}}{\alpha}\right)^{-\alpha}$

The responses are positive. The location vector contains $\mu_j > 0$, the scale vector contains $\sigma_j > 0$, and the parameter α is also positive. It includes the multivariate logistic distribution ($\alpha = 1$). The marginal cdf is a three parameter generalized logistic distribution, $F(y) = \left(1+\frac{\mathrm{e}^{-\frac{y-\mu}{\sigma}}}{\alpha}\right)^{-\alpha}$, corresponding to Equation (B.32) where $\alpha = \frac{1}{\sigma}$ (in both).
(Cook and Johnson, 1981)

One parameter distribution

- Density $f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)\alpha^{\alpha}}{\Gamma(\alpha)} \left(\alpha + \sum_{j=1}^{m} \mathrm{e}^{-y_j}\right)^{-\alpha-m} \prod_{j=1}^{m} \mathrm{e}^{-y_j}$
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1 + \frac{\sum_{j=1}^{m} \mathrm{e}^{-y_j}}{\alpha}\right)^{-\alpha}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the parameter α is also positive.
(Cook and Johnson, 1981)

- Density $f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)}{\Gamma(\alpha)} \left(1 + \sum_{j=1}^{m} \mathrm{e}^{-y_j}\right)^{-\alpha-m} \prod_{j=1}^{m} \mathrm{e}^{-y_j}$ (B.53)
- CDF $F_{\mathrm{M}}(\mathbf{y}) = \left(1 + \sum_{j=1}^{m} \mathrm{e}^{-y_j}\right)^{-\alpha}$

The responses are positive. The location is set to zero ($\mu = 0$), the scale is set to unity ($\sigma = 1$), and the parameter α is also positive. This form is obtained by transforming the response $\mathrm{e}^{-z_j} = \alpha \mathrm{e}^{-y_j}$ with Jacobian α.
(Cook and Johnson, 1981)

MULTIVARIATE PARETO DISTRIBUTION

Two parameter distribution

- Density

$$f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)}{\Gamma(\alpha)\prod_{j=1}^{m}\theta_j} \left(\sum_{j=1}^{m} \frac{y_j}{\theta_j} - m + 1\right)^{-\alpha-m} \qquad \text{(B.54)}$$

- Survival function $S_{\mathrm{M}}(\mathbf{y}) = \left(\sum_{j=1}^{m} \frac{y_j}{\theta_j} - m + 1\right)^{-\alpha}$

The responses are positive, the parameter α is positive, and $0 < \theta_j < y_j$. The marginal cdf is a two parameter Pareto distribution, $F(y) = 1 - \frac{\theta}{y^{\alpha}}$, corresponding to Equation (B.25) where $\theta = \kappa^{\alpha}$.
(Johnson and Kotz, 1972, p. 286; Cook and Johnson, 1981 [method corrected])

MULTIVARIATE EXPONENTIAL DISTRIBUTION

One parameter distribution

- Density

$$f_{\mathrm{M}}(\mathbf{y}) = \frac{\Gamma(\alpha+m)}{\Gamma(\alpha)\alpha^{m}} \left(\sum_{j=1}^{m} \mathrm{e}^{\frac{y_j}{\alpha}} - m + 1\right)^{-\alpha-m} \prod_{j=1}^{m} \mathrm{e}^{\frac{y_j}{\alpha}} \qquad \text{(B.55)}$$

- Survival function $S_{\mathrm{M}}(\mathbf{y}) = \left(\sum_{j=1}^{m} \mathrm{e}^{\frac{y_j}{\alpha}} - m + 1\right)^{-\alpha}$

The responses are positive and the parameter α is positive. The marginal cdf is a unit exponential distribution, $F(y) = 1 - \mathrm{e}^{-y}$, corresponding to Equation (B.7).
(Johnson and Kotz, 1972, p. 288 [corrected])

MULTIVARIATE POWER EXPONENTIAL DISTRIBUTION

Three parameter distribution

- Density $f_{\mathrm{M}}(\mathbf{y}) = \frac{m\Gamma\left(\frac{m}{2}\right)\mathrm{e}^{-\frac{1}{2}\left[(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{\kappa}}}{2^{1+\frac{m}{2\kappa}}\pi^{\frac{m}{2}}\sqrt{|\boldsymbol{\Sigma}|}\,\Gamma\left(1+\frac{m}{2\kappa}\right)}$ (B.56)

The responses vector is $\mathbf{y}$. The location vector is $\boldsymbol{\mu}$, the scale is the definite

positive symmetric matrix $\mathbf{\Sigma}$, and the family is $\kappa > 0$. It includes the multivariate Laplace distribution ($\kappa = 0.5$), the multivariate Gaussian distribution ($\kappa = 1$), and the multivariate uniform distribution ($\kappa \to \infty$).
(Gómez, E., Gómez-Villegas, M.A., and Marin, J.M., 1998)

MULTIVARIATE STUDENT T DISTRIBUTION

Three parameter distribution

- Density $\quad f_{\mathrm{M}}(\mathbf{y}) = \dfrac{\Gamma\left(\frac{\kappa+m}{2}\right)}{(\kappa\pi)^{\frac{m}{2}}\sqrt{|\mathbf{\Sigma}|}\Gamma\left(\frac{\kappa}{2}\right)\left[1+\frac{1}{\kappa}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{\frac{\kappa+m}{2}}} \qquad \text{(B.57)}$

The responses vector is $\mathbf{y}$ and the parameter κ is the number of degrees of freedom. The location vector is $\boldsymbol{\mu}$ and the scale is the definite positive symmetric matrix $\mathbf{\Sigma}$. It includes the multivariate Cauchy distribution ($\kappa = 1$) and multivariate Gaussian distribution ($\kappa \to \infty$).
(Johnson and Kotz, 1972, p. 134)

MULTIVARIATE GAUSSIAN DISTRIBUTION

Two parameter distribution

- Density $\quad f_{\mathrm{M}}(\mathbf{y}) = \dfrac{\mathrm{e}^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}}{(2\pi)^{\frac{m}{2}}\sqrt{|\mathbf{\Sigma}|}} \qquad \text{(B.58)}$

The responses vector is $\mathbf{y}$. The location vector is $\boldsymbol{\mu}$ and the scale is the definite positive symmetric matrix $\mathbf{\Sigma}$.
(Johnson and Kotz, 1972, pp. 37–38)

Appendix C
Archimedean copulas

C.1 One dependence parameter

LAPLACE TRANSFORMS

(1) The gamma Laplace transform

$$\phi(s) = (1+\delta s)^{-1/\delta} \qquad \text{or} \qquad \phi(s) = (1+s)^{-1/\delta}$$

and its inverse

$$\phi^{-1}(t) = \frac{t^{-\delta}-1}{\delta} \qquad \text{or} \qquad \phi^{-1}(t) = t^{-\delta}-1$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1,u_2) = (u^{-\delta}+v^{-\delta}-1)^{-1/\delta}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left(\sum_{j=1}^{m} u_j^{-\delta} - (m-1)\right)^{-1/\delta}$$

where $\delta \geq 0$.
Independence and the Fréchet upper bound respectively occur at values 0 and ∞ of the dependence parameter δ.
Figure C.1 shows the bivariate dependence for $\delta = 2$ with uniform, Gaussian, lognormal, gamma ($\alpha = 2$), logistic, Cauchy, non-central beta ($\alpha = 2$, $\beta = 4$, and non-central parameter equals 3), Laplace, and power exponential ($\mu = 0$, $\sigma = 1$, and $\kappa = 3$) margins.
This copula is sometimes called the Clayton, generalized Cook and Johnson, or the Pareto family (Kimeldorf and Sampson, 1975; Clayton, 1978; Cook and Johnson, 1983).

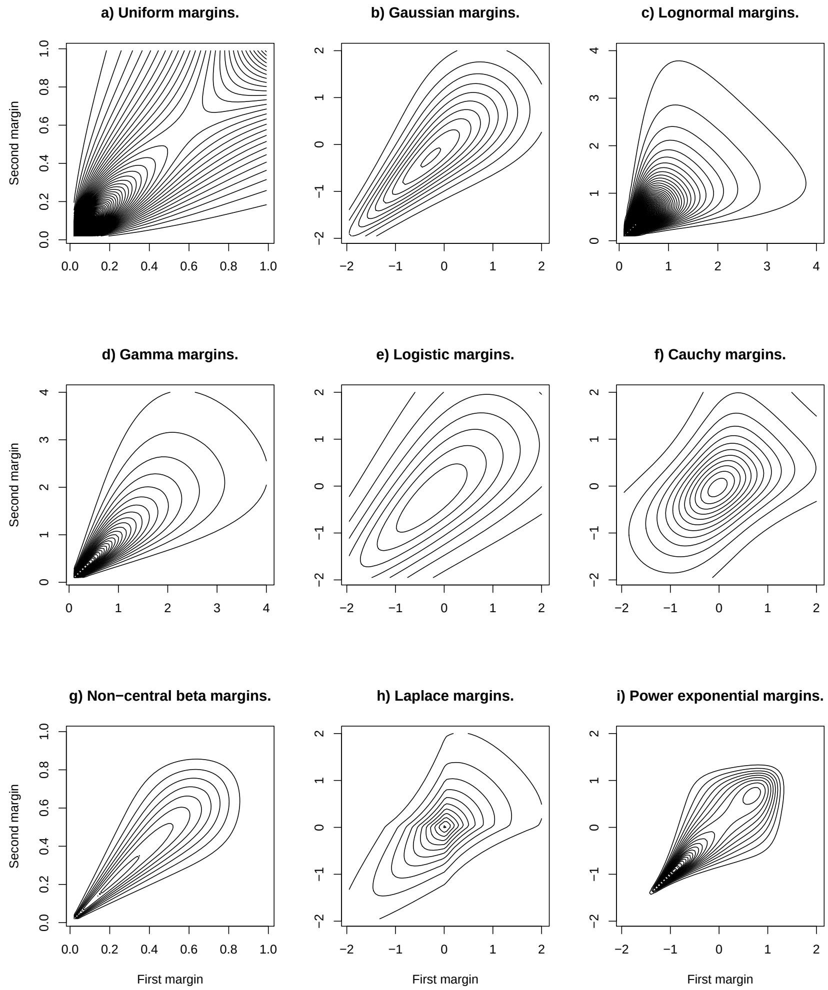

Fig. C.1. Density contours for a bivariate Archimedean copula generated from a gamma Laplace transform with a) uniform, b) Gaussian, c) lognormal, d) gamma, e) logistic, f) Cauchy, g) non-central beta, h) Laplace, and i) power exponential margins ($\delta = 2$).

(2) The positive stable Laplace transform

$$\phi(s) = \mathrm{e}^{(-s^{1/\delta})}$$

and its inverse

$$\phi^{-1}(t) = (-\ln t)^{\delta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1,u_2) = \mathrm{e}^{-\left[(-\ln u)^{\delta}+(-\ln v)^{\delta}\right]^{1/\delta}}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{-\left[\sum_{j=1}^{m}(-\ln u_j)^{\delta}\right]^{1/\delta}}$$

where $\delta \geq 1$.
Independence and the Fréchet upper bound respectively occur at values 1 and ∞ of the dependence parameter δ.
Figure C.2 shows the bivariate dependence for $\delta = 3.5$ with uniform, non-central beta ($\alpha = 2$, $\beta = 4$, and non-central parameter equals 3), and Levy margins.
This copula is sometimes called the Gumbel family (Gumbel, 1960b).

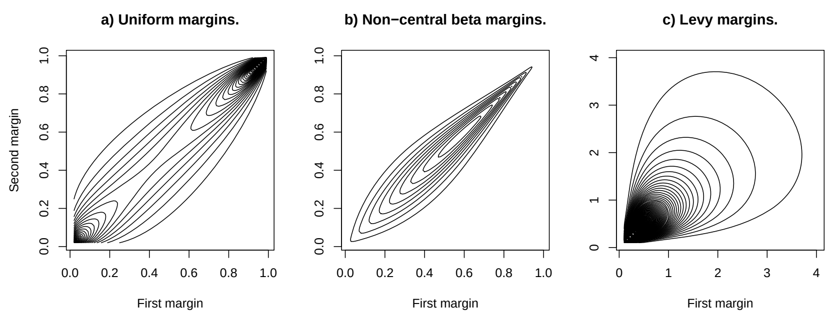

Fig. C.2. Density contours for a bivariate Archimedean copula generated from a positive stable Laplace transform with a) uniform, b) non-central beta, and c) Levy margins ($\delta = 3.5$).

(3) The logarithmic series Laplace transform

$$\phi(s) = -\frac{1}{\delta}\ln\left[1 + \mathrm{e}^{-s}\left(\mathrm{e}^{-\delta} - 1\right)\right]$$

and its inverse

$$\phi^{-1}(t) = -\ln\frac{\mathrm{e}^{-\delta t} - 1}{\mathrm{e}^{-\delta} - 1}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = -\frac{1}{\delta}\ln\left[1 + \frac{(\mathrm{e}^{-\delta u} - 1)(\mathrm{e}^{-\delta v} - 1)}{\mathrm{e}^{-\delta} - 1}\right]$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = -\frac{1}{\delta}\ln\left[1 + \frac{\prod_{j=1}^{m}(\mathrm{e}^{-\delta u_j} - 1)}{(\mathrm{e}^{-\delta} - 1)^{m-1}}\right]$$

where $\delta \geq 0$.
Independence and the Fréchet upper bound respectively occur at values 0 and ∞ of the dependence parameter δ.
Figure C.3 shows the bivariate dependence for $\delta = 5$ with uniform, logistic, and Burr ($\mu = 2$, $\sigma = 2$, And $\kappa = 2$) margins.
This copula is sometimes called the Frank family (Frank, 1979).

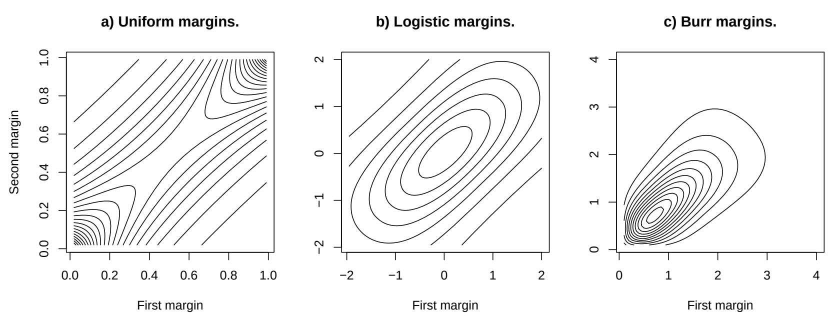

Fig. C.3. Density contours for a bivariate Archimedean copula generated from a logarithmic series Laplace transform with a) uniform, b) logistic, and c) Burr margins ($\delta = 5$).

(4) The power series Laplace transform

$$\phi(s) = 1 - \left(1 - \mathrm{e}^{-s}\right)^{1/\delta}$$

and its inverse

$$\phi^{-1}(t) = -\ln\left[1 - (1-t)^{\delta}\right]$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = 1 - \left[(1-u)^{\delta} + (1-v)^{\delta} - (1-u)^{\delta}(1-v)^{\delta}\right]^{1/\delta}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = 1 - \left\{1 - \prod_{j=1}^{m}\left[1 - (1-u_j)^{\delta}\right]\right\}^{1/\delta}$$

where $\delta \geq 1$.
Independence and the Fréchet upper bound respectively occur at values 1 and ∞ of the dependence parameter δ.
Figure C.4 shows the bivariate dependence for $\delta = 3$ with uniform, gamma ($\alpha = 2$), and Cauchy margins.
This copula is sometimes called the Joe family (Joe, 1993).

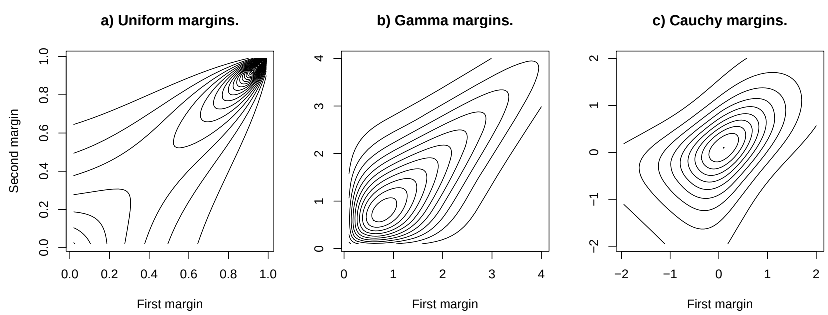

Fig. C.4. Density contours for a bivariate Archimedean copula generated from a power series Laplace transform with a) uniform, b) gamma, and c) Cauchy margins ($\delta = 3$).

(5) The inverse Gaussian Laplace transform

$$\phi(s) = \mathrm{e}^{\frac{1}{\delta} - \sqrt{\frac{1}{\delta}\left(\frac{1}{\delta} + 2s\right)}}$$

and its inverse

$$\phi^{-1}(t) = \frac{\delta\left(\frac{1}{\delta} - \ln[t]\right)^2 - \frac{1}{\delta}}{2}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \mathrm{e}^{\frac{1}{\delta} - \sqrt{\left(\frac{1}{\delta} - \ln[u]\right)^2 + \left(\frac{1}{\delta} - \ln[v]\right)^2 - \frac{1}{\delta^2}}}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{\frac{1}{\delta} - \sqrt{\Sigma_{j=1}^{m}\left(\frac{1}{\delta} - \ln[u_j]\right)^2 - (m-1)\frac{1}{\delta^2}}}$$

where $\delta \geq 0$.
Independence and the Fréchet upper bound respectively occur at values 0 and ∞ of the dependence parameter δ.
Figure C.5 shows the bivariate dependence for $\delta = 10$ with uniform, inverse Gaussian ($\mu = 2$ and $\sigma = 1$), and power exponential ($\mu = 0$, $\sigma = 1$, and $\kappa = 3$) margins.

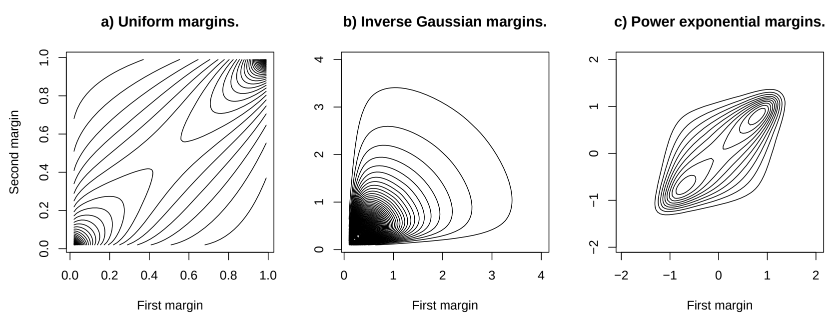

Fig. C.5. Density contours for a bivariate Archimedean copula generated from an inverse Gaussian Laplace transform with a) uniform, b) inverse Gaussian, and c) power exponential margins ($\delta = 10$).

DECREASING CONVEX GENERATORS

(1) The generator

$$\phi^{-1}(t) = \ln \frac{1-\theta(1-t)}{t} \qquad \text{or} \qquad \phi^{-1}(t) = \frac{1}{1-\theta} \ln \frac{1+\theta(t-1)}{t}$$

and its inverse

$$\phi(s) = \frac{1-\theta}{\mathrm{e}^{s}-\theta} \qquad \text{or} \qquad \phi(s) = \frac{\theta-1}{\theta-\mathrm{e}^{(1-\theta)s}}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1,u_2) = \frac{uv}{1-\delta(1-u)(1-v)}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \frac{\delta-1}{\delta-\prod_{j=1}^{m}\left[\frac{1+\delta(u_j-1)}{u_j}\right]}$$

where $-1 < \delta < 1$. Independence occurs at value 0 of the dependence parameter δ. Figure C.6 shows the bivariate dependence for $\delta = 0.5$ with uniform, Burr ($\mu = 2$, $\sigma = 2$, and $\kappa = 2$), and simplex ($\mu = 0.5$ and $\sigma = 1$) margins. This copula is sometimes called the Ali-Mikhail-Haq family (Ali *et al.*, 1978).

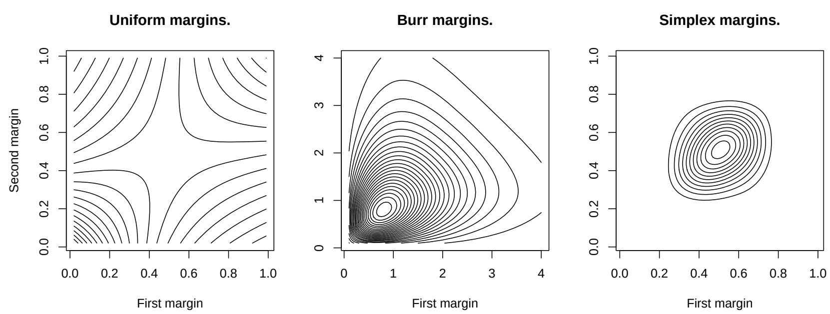

Fig. C.6. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Burr, and c) simplex margins ($\delta = 0.5$).

(2) The generator

$$\phi^{-1}(t) = \frac{1-t}{1+(\theta-1)t}$$

and its inverse

$$\phi(s) = \frac{1-s}{1+(\theta-1)s}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max\left(\frac{\delta^2 uv - (1-u)(1-v)}{\delta^2 - (\delta-1)^2(1-u)(1-v)}, 0\right)$$

where $\delta > 1$. The Fréchet lower bound occurs at value 1 of the dependence parameter δ. Figure C.7 shows the bivariate dependence for $\delta = 100$ with uniform, Gaussian, and Cauchy margins.

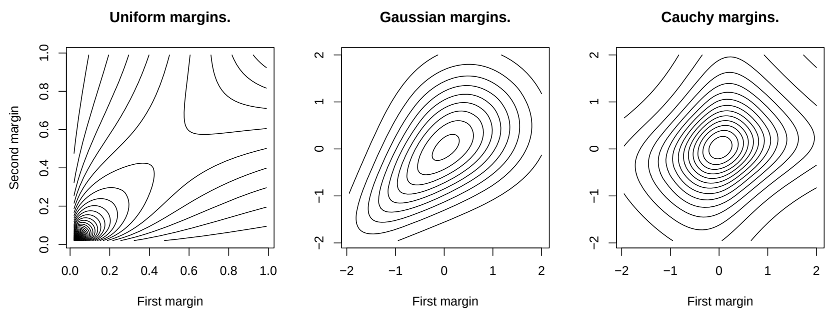

Fig. C.7. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Gaussian, and c) Cauchy margins ($\delta = 100$).

(3) The generator

$$\phi^{-1}(t) = \ln(1 - \theta \ln t)$$

and its inverse

$$\phi(s) = \mathrm{e}^{\frac{1-\mathrm{e}^s}{\theta}}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = uv\mathrm{e}^{-\delta \ln u \ln v}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{\frac{1-\prod_{j=1}^{m}[1-\delta \ln u_j]}{\delta}}$$

where $0 \leq \delta < 1$. Independence occurs at value 0 of the dependence parameter δ. Figure C.8 shows the bivariate dependence for $\delta = 0.8$ with uniform, gamma ($\alpha = 2$), and non-central beta ($\alpha = 2$, $\beta = 4$, and non-central parameter equals 3) margins. This copula is sometimes called the Gumbel-Barnett family (Gumbel, 1960a; Barnett, 1980).

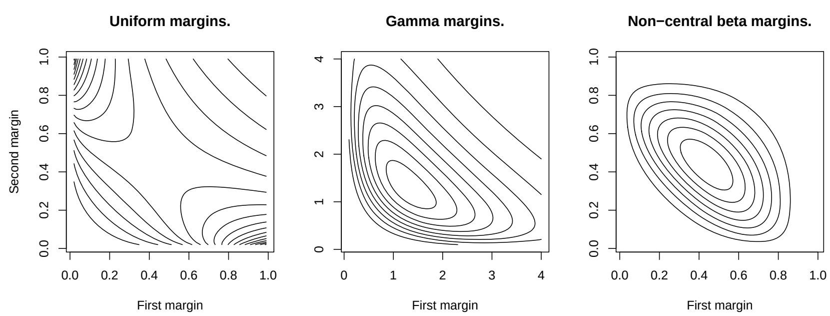

Fig. C.8. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) gamma, and c) non-central beta margins ($\delta = 0.8$).

(4) The generator

$$\phi^{-1}(t) = \ln\left(2t^{-\theta} - 1\right)$$

and its inverse

$$\phi(s) = \left(\frac{1+\mathrm{e}^{s}}{2}\right)^{-1/\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \frac{uv}{\left[1 + (1-u^{\delta})(1-v^{\delta})\right]^{1/\delta}}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left[\frac{1 + \prod_{j=1}^{m}\left(2u_j^{-\delta} - 1\right)}{2}\right]^{-1/\delta}$$

where $0 \le \delta < 1$. Independence occurs at value 0 of the dependence parameter δ. Figure C.9 shows the bivariate dependence for $\delta = 0.7$ with uniform, logistic, and power exponential ($\mu = 0$, $\sigma = 1$, and $\kappa = 3$) margins.

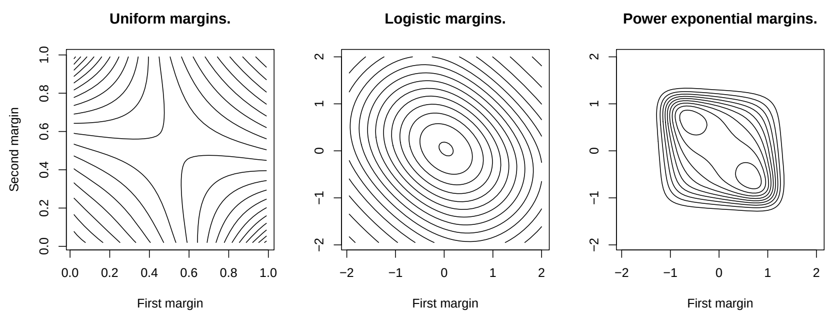

Fig. C.9. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) logistic, and c) power exponential margins ($\delta = 0.7$).

(5) The generator

$$\phi^{-1}(t) = \ln\left(2 - t^{\theta}\right)$$

and its inverse

$$\phi(s) = (2 - \mathrm{e}^{s})^{1/\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max\left(\left[(uv)^{\delta} - 2\left(1 - u^{\delta}\right)\left(1 - v^{\delta}\right)\right]^{1/\delta}, 0\right)$$

where $0 \le \delta < \frac{1}{2}$. Independence occurs at value 0 of the dependence parameter δ. Figure C.10 shows the bivariate dependence for $\delta = 0.2$ with uniform, Gaussian, and lognormal margins.

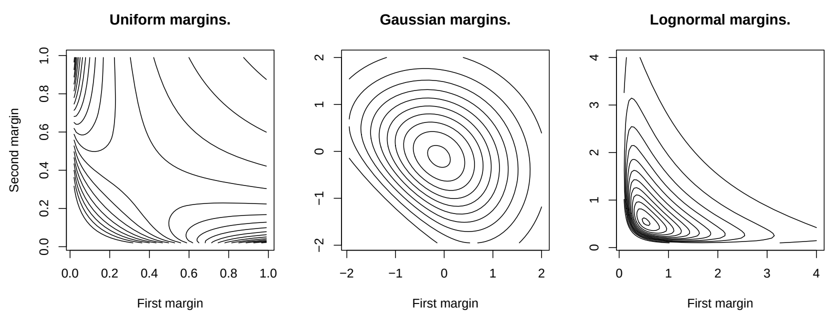

Fig. C.10. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Gaussian, and c) lognormal margins ($\delta = 0.2$).

(6) The generator

$$\phi^{-1}(t) = \left(\frac{1}{t} - 1\right)^{\theta}$$

and its inverse

$$\phi(s) = \frac{1}{1 + s^{1/\theta}}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \frac{1}{1 + \left[\left(\frac{1}{u} - 1\right)^{\delta} + \left(\frac{1}{v} - 1\right)^{\delta}\right]^{1/\delta}}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \frac{1}{1 + \left[\sum_{j=1}^{m} \left(\frac{1}{u_j} - 1\right)^{\delta}\right]^{1/\delta}}$$

where $\delta > 1$. The Fréchet upper bound occurs at value ∞ of the dependence parameter δ. Figure C.11 shows the bivariate dependence for $\delta = 2$ with uniform, Laplace, and Pareto ($\mu = 2$ and $\sigma = 2$) margins.

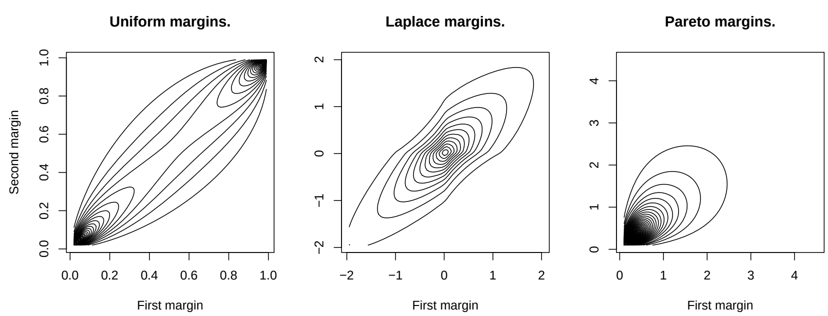

Fig. C.11. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Laplace, and c) Pareto margins ($\delta = 2$).

(7) The generator

$$\phi^{-1}(t) = (1 - \ln t)^{\theta} - 1$$

and its inverse

$$\phi(s) = \mathrm{e}^{1-(1+s)^{1/\theta}}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \mathrm{e}^{1-\left[(1-\ln u)^{\delta}+(1-\ln v)^{\delta}-1\right]^{1/\delta}}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{1-\left[\sum_{j=1}^{m}(1-\ln u_j)^{\delta}-(m-1)\right]^{1/\delta}}$$

where $\delta > 0$. Independence and the Fréchet upper bound respectively occur at values 1 and ∞ of the dependence parameter δ. Figure C.12 shows the bivariate dependence for $\delta = 7$ with uniform, Cauchy, and Levy margins.

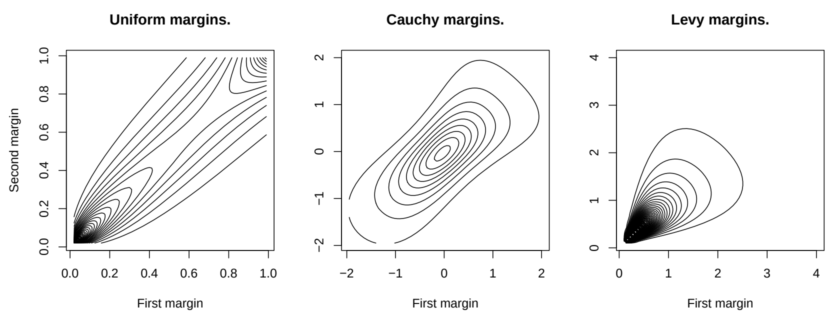

Fig. C.12. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Cauchy, and c) Levy margins ($\delta = 7$).

(8) The generator

$$\phi^{-1}(t) = \left(t^{-1/\theta} - 1\right)^{\theta}$$

and its the inverse

$$\phi(s) = \left(1 + s^{1/\theta}\right)^{-\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \left\{1 + \left[\left(u^{-1/\delta} - 1\right)^{\delta} + \left(v^{-1/\delta} - 1\right)^{\delta}\right]^{1/\delta}\right\}^{-\delta}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left\{1 + \left[\sum_{j=1}^{m}\left(u_j^{-1/\delta} - 1\right)^{\delta}\right]^{1/\delta}\right\}^{-\delta}$$

where $\delta > 1$. The Fréchet upper bound occurs at value ∞ of the dependence parameter δ. Figure C.13 shows the bivariate dependence for $\delta = 2$ with uniform, Gaussian, and inverse Gaussian ($\mu = 2$ and $\sigma = 1$) margins.

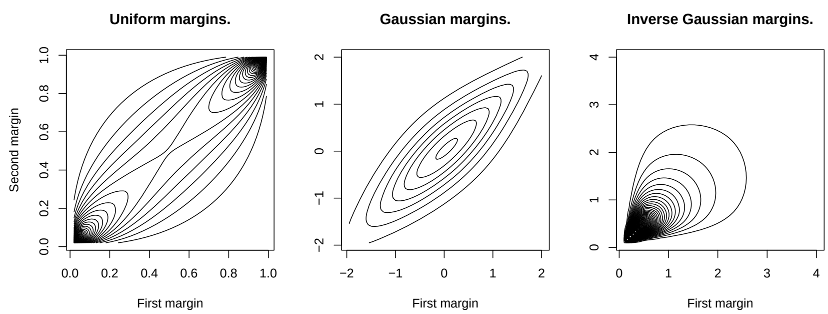

Fig. C.13. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Gaussian, and c) inverse Gaussian margins ($\delta = 2$).

(9) The generator

$$\phi^{-1}(t) = \left(1 - t^{1/\theta}\right)^{\theta}$$

and its inverse

$$\phi(s) = \left(1 - s^{1/\theta}\right)^{\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max\left(\left\{1 - \left[\left(1 - u^{1/\delta}\right)^{\delta} + \left(1 - v^{1/\delta}\right)^{\delta}\right]^{1/\delta}\right\}^{\delta}, 0\right)$$

where $\delta > 1$. The Fréchet lower and upper bounds respectively occur at values 1 and ∞ of the dependence parameter δ. Figure C.14 shows the bivariate dependence for $\delta = 3$ with uniform, Gamma ($\alpha = 2$), and Pareto ($\mu = 2$ and $\sigma = 2$) margins. This copula is sometimes called the Genest-Ghoudi family (Genest and Ghoudi, 1994).

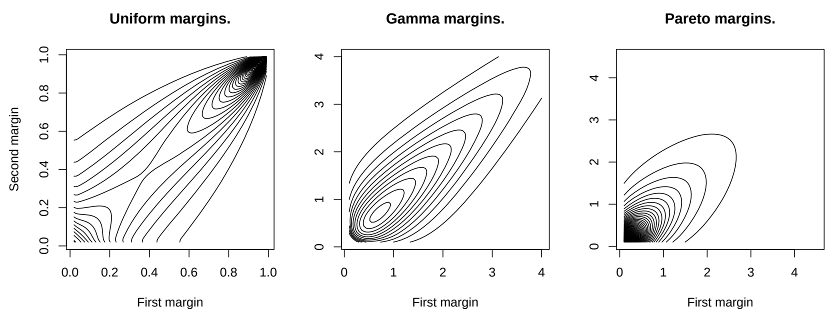

Fig. C.14. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Gamma, and c) Pareto margins ($\delta = 3$).

(10) The generator

$$\phi^{-1}(t) = \left(1 + \frac{\theta}{t}\right)(1-t)$$

and its inverse

$$\phi(s) = \frac{1 - s - \theta + \sqrt{(s+\theta-1)^2 + 4\theta}}{2}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \frac{1}{2}\left\{u + v - 1 - \delta\left(\frac{1}{u} + \frac{1}{v} - 1\right)\right. \\ \left. + \sqrt{\left[u + v - 1 - \delta\left(\frac{1}{u} + \frac{1}{v} - 1\right)\right]^2 + 4\delta}\right\}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \frac{1 - \left[\sum_{j=1}^{m}\left(1 + \frac{\delta}{u_j}\right)(1-u_j)\right] - \delta}{2} \\ + \frac{\sqrt{\left\{\left[\sum_{j=1}^{m}\left(1 + \frac{\delta}{u_j}\right)(1-u_j)\right] + \delta - 1\right\}^2 + 4\delta}}{2}$$

where $\delta > 0$. The Fréchet lower bound occurs at value 0 of the dependence parameter δ. Figure C.15 shows the bivariate dependence for $\delta = 50$ with uniform, non-central beta ($\alpha = 2$, $\beta = 4$, and non-central parameter equals 3), and Burr ($\mu = 2$, $\sigma = 2$, and $\kappa = 2$) margins.

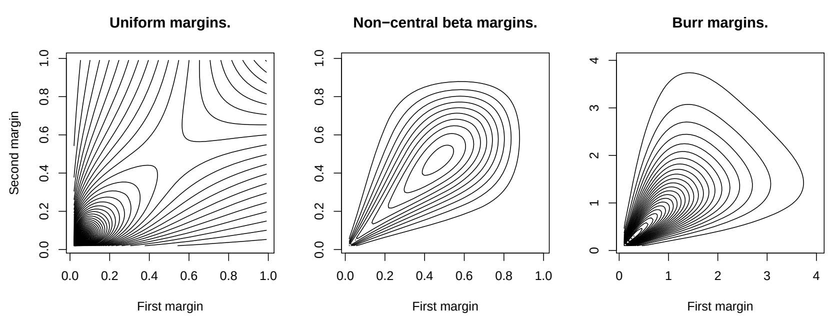

Fig. C.15. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) non-central beta, and c) Burr margins ($\delta = 50$).

(11) The generator

$$\phi^{-1}(t) = -\ln\frac{(1+t)^{-\theta}-1}{2^{-\theta}-1}$$

and its inverse

$$\phi(s) = \left[1+(2^{-\theta}-1)\mathrm{e}^{-s}\right]^{-1/\theta}-1$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1,u_2) = \left\{1+\frac{[(1+u)^{-\delta}-1][(1+v)^{-\delta}-1]}{2^{-\delta}-1}\right\}^{-1/\delta}-1$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left\{1+\frac{\prod_{j=1}^{m}[(1+u_j)^{-\delta}-1]}{(2^{-\delta}-1)^{m-1}}\right\}^{-1/\delta}-1$$

where $\delta \neq 0$. Independence and the Fréchet upper bound respectively occur at values -1 and ∞ of the dependence parameter δ. Figure C.16 shows the bivariate dependence for $\delta = -10$ with uniform, inverse Gaussian ($\mu = 2$ and $\sigma = 1$), and simplex ($\mu = 0.5$ and $\sigma = 1$) margins.

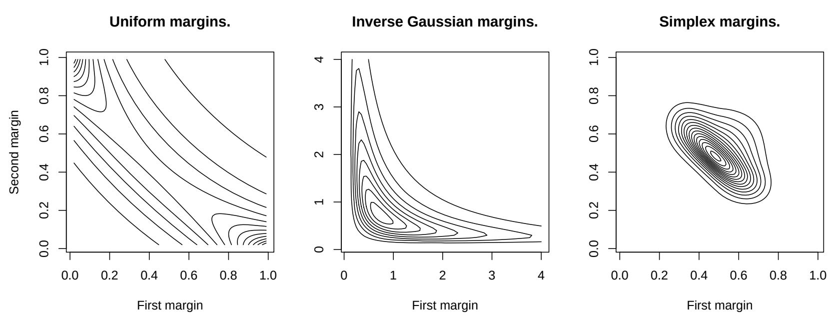

Fig. C.16. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) inverse Gaussian, and c) simplex margins ($\delta = -10$).

(12) The generator

$$\phi^{-1}(t) = \mathrm{e}^{\frac{\theta}{t}} - \mathrm{e}^{\theta}$$

and its inverse

$$\phi(s) = \frac{\theta}{\ln\left(s + \mathrm{e}^{\theta}\right)}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \frac{\delta}{\ln\left(\mathrm{e}^{\frac{\delta}{u}} + \mathrm{e}^{\frac{\delta}{v}} - \mathrm{e}^{\delta}\right)}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \frac{\delta}{\ln\left(\sum_{j=1}^{m} \mathrm{e}^{\frac{\delta}{u_j}} - (m-1)\mathrm{e}^{\delta}\right)}$$

where $\delta > 0$. The Fréchet upper bound occurs at value ∞ of the dependence parameter δ. Figure C.17 shows the bivariate dependence for $\delta = 1$ with uniform, logistic, and non-central beta ($\alpha = 2$, $\beta = 4$, and non-central parameter equals 3) margins.

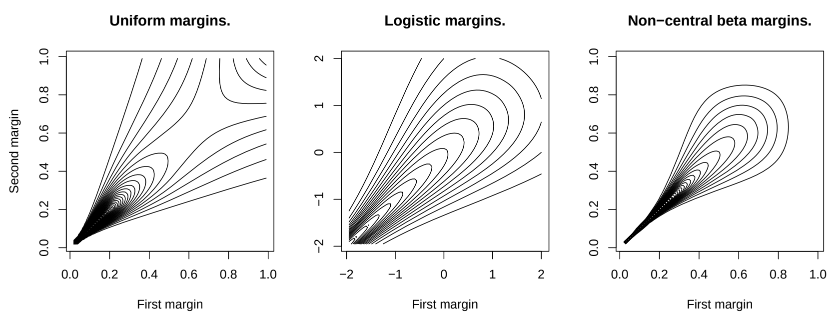

Fig. C.17. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) logistic, and c) non-central beta margins ($\delta = 1$).

(13) The generator

$$\phi^{-1}(t) = \mathrm{e}^{t^{-\theta}} - \mathrm{e}$$

and its inverse

$$\phi(s) = [\ln(s+\mathrm{e})]^{-1/\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1,u_2) = \left[\ln\left(\mathrm{e}^{u^{-\delta}} + \mathrm{e}^{v^{-\delta}} - \mathrm{e}\right)\right]^{-1/\delta}$$

and the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left[\ln\left(\sum_{j=1}^{m} \mathrm{e}^{u_j^{-\delta}} - (m-1)\mathrm{e}\right)\right]^{-1/\delta}$$

where $\delta \geq 0$. Independence and the Fréchet upper bound respectively occur at values 0 and ∞ of the dependence parameter δ. Figure C.18 shows the bivariate dependence for $\delta = 0.5$ with uniform, Levy, and power exponential margins.

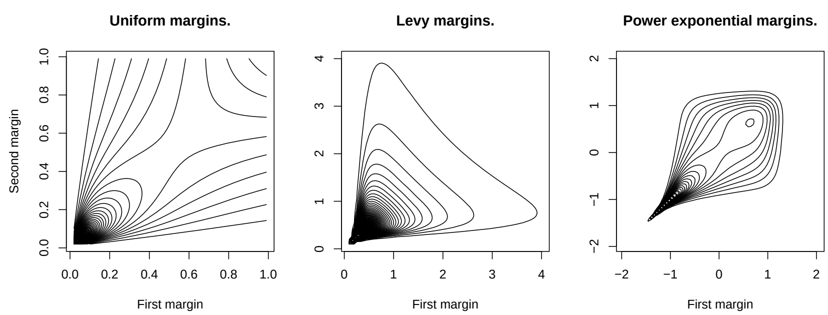

Fig. C.18. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) Levy, and c) power exponential margins ($\delta = 0.5$).

(14) The generator

$$\phi^{-1}(t) = 1 - \left[1 - (1-t)^{\theta}\right]^{1/\theta}$$

and its inverse

$$\phi(s) = 1 - \left[1 - (1-s)^{\theta}\right]^{1/\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = 1 - \left[1 - \left\{\max\left(\left[1-(1-u)^{\delta}\right]^{1/\delta} + \left[1-(1-v)^{\delta}\right]^{1/\delta} - 1, 0\right)\right\}^{\delta}\right]^{1/\delta}$$

where $\delta > 1$. The Fréchet lower and upper bounds respectively occur at values 1 and ∞ of the dependence parameter δ. Figure C.19 shows the bivariate dependence for $\delta = 3$ with uniform, lognormal, and Laplace margins.

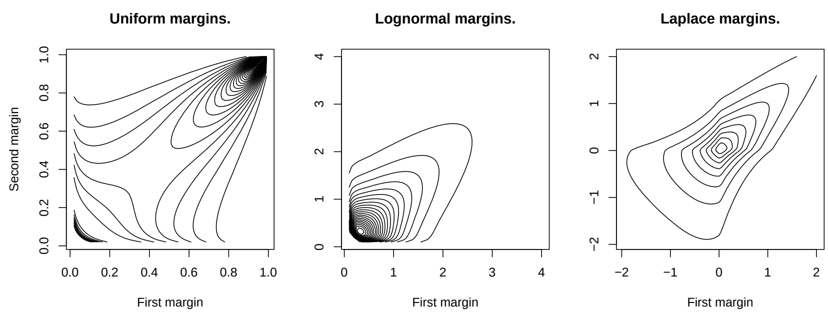

Fig. C.19. Density contours for a bivariate Archimedean copula generated from a decreasing convex generators with a) uniform, b) lognormal, and c) Laplace margins ($\delta = 3$).

(15) The generator

$$\phi^{-1}(t) = (1-t)^{\theta}$$

and its inverse

$$\phi(s) = 1 - s^{1/\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max\left(1 - \left[(1-u)^{\delta} + (1-v)^{\delta}\right]^{1/\delta}, 0\right)$$

where $\delta > 1$. The Fréchet lower and upper bounds respectively occur at values 1 and ∞ of the dependence parameter δ.

(16) The generator

$$\phi^{-1}(t) = -\ln[\theta t + (1-\theta)]$$

and its inverse

$$\phi(s) = \frac{\mathrm{e}^{-s} - (1-\theta)}{\theta}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max(\delta uv + (1-\delta)(u+v-1), 0)$$

where $0 < \delta \leq 1$. The Fréchet lower bound and independence respectively occur at values 0 and 1 of the dependence parameter δ.

(17) The generator

$$\phi^{-1}(t) = \mathrm{e}^{\frac{\theta}{t-1}}$$

and its inverse

$$\phi(s) = 1 + \frac{\theta}{\ln s}$$

generate the bivariate Archimedean copula

$$C_{\mathrm{B}}(u_1, u_2) = \max\left(1 + \frac{\delta}{\ln\left[\mathrm{e}^{\frac{\delta}{u-1}} + \mathrm{e}^{\frac{\delta}{v-1}}\right]}, 0\right)$$

where $\delta > 2$. The Fréchet upper bound occurs at value ∞ of the dependence parameter δ.

C.2 Two dependence parameters

(1) Applying Equation (4.4) to a positive stable (ψ) and a gamma (ϕ) Laplace transform, the Laplace transform obtained is

$$\phi(s) = \left(1 + s^{\frac{1}{\delta}}\right)^{-\frac{1}{\theta}}$$

with inverse

$$\phi^{-1}(t) = \left(t^{-\theta} - 1\right)^{\delta}$$

which can also be obtained by applying Equation (4.3) to the generator

$$\psi(t) = \frac{1}{t} - 1$$

The multivariate Archimedean copula is

$$C_{\mathrm{M}}(\mathbf{u}) = \left(1 + \left[\sum_{j=1}^{m} \left(u_j^{-\theta} - 1\right)^{\delta}\right]^{\frac{1}{\delta}}\right)^{-\frac{1}{\theta}}$$

where $\delta \geq 1$ and $\theta > 0$.
(Joe, 1997, p. 150; Nelsen, 1999, p. 116)

(2) Applying Equation (4.4) to two gamma Laplace transform, the Laplace transform obtained is

$$\phi(s) = \left(1 + \frac{\ln[1+s]}{\delta}\right)^{-\frac{1}{\theta}}$$

with inverse

$$\phi^{-1}(t) = \mathrm{e}^{\delta\left(t^{-\theta}-1\right)} - 1$$

The multivariate Archimedean copula is

$$C_{\mathrm{M}}(\mathbf{u}) = \left(1 + \frac{1}{\delta}\ln\left[\sum_{j=1}^{m} \mathrm{e}^{\delta\left(u_j^{-\theta}-1\right)} - (m-1)\right]\right)^{-\frac{1}{\theta}}$$

where $\delta > 0$ and $\theta > 0$.
(Joe, 1997, pp. 150–151)

(3) Applying Equation (4.4) to a gamma (ψ) and a positive stable (ϕ) Laplace transform, the Laplace transform obtained is

$$\phi(s) = \mathrm{e}^{-\left(\frac{\ln[1+s]}{\delta}\right)^{\frac{1}{\theta}}}$$

with inverse

$$\phi^{-1}(t) = \mathrm{e}^{\delta(-\ln[t])^{\theta}} - 1$$

The multivariate Archimedean copula is

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{-\left(\frac{1}{\delta}\ln\left[\Sigma_{j=1}^{m}\mathrm{e}^{\delta(-\ln[u_j])^{\theta}} - (m-1)\right]\right)^{\frac{1}{\theta}}}$$

where $\delta > 0$ and $\theta \geq 1$.
(Joe, 1997, p. 151)

(4) Applying Equation (4.4) to a positive stable (ψ) and a power series (ϕ) Laplace transform, the Laplace transform obtained is

$$\phi(s) = 1 - \left[1 - \mathrm{e}^{-s^{\frac{1}{\delta}}}\right]^{\frac{1}{\theta}}$$

with inverse

$$\phi^{-1}(t) = \left(-\ln\left[1-(1-t)^{\theta}\right]\right)^{\delta}$$

The multivariate Archimedean copula is

$$C_{\mathrm{M}}(\mathbf{u}) = 1 - \left(1 - \mathrm{e}^{-\left[\Sigma_{j=1}^{m}\left(-\ln\left[1-(1-u_j)^{\theta}\right]\right)^{\delta}\right]^{\frac{1}{\delta}}}\right)^{\frac{1}{\theta}}$$

where $\delta \geq 1$ and $\theta \geq 1$.
(Joe, 1997, pp. 152–153)

(5) Applying Equation (4.4) to a gamma (ψ) and a power series (ϕ) Laplace transform, the Laplace transform obtained is

$$\phi(s) = 1 - \left[1 - (1+s)^{-\frac{1}{\delta}}\right]^{\frac{1}{\theta}}$$

with inverse

$$\phi^{-1}(t) = \left[1-(1-t)^{\theta}\right]^{-\delta} - 1$$

The multivariate Archimedean copula is

$$C_{\mathrm{M}}(\mathbf{u}) = 1 - \left[1 - \left(\sum_{j=1}^{m}\left[1-(1+u_j)^{\theta}\right]^{-\delta} - (m-1)\right)^{-\frac{1}{\delta}}\right]^{\frac{1}{\theta}}$$

where $\delta > 0$ and $\theta \geq 1$.
(Joe, 1997, p. 153)

(6) The Laplace transform of a discrete power series, on positive integers with mass $\pi_i = \frac{[1-(1-\delta)^\theta]^i \prod_{j=1}^{i-1}(j-\frac{1}{\theta})}{\delta\theta i!}$ for $i > 1$ and $\pi_1 = \frac{1-(1-\delta)^\theta}{\theta\delta}$,

$$\phi(s) = \frac{1 - \left[1 - \left(1 - [1-\delta]^\theta\right) \mathrm{e}^{-s}\right]^{\frac{1}{\theta}}}{\delta}$$

and its inverse

$$\phi^{-1}(t) = -\ln\left[\frac{1-(1-\delta t)^\theta}{1-(1-\delta)^\theta}\right]$$

generate the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \frac{1}{\delta}\left[1 - \left(1 - \frac{\prod_{j=1}^m \left[1-(1-\delta u_j)^\theta\right]}{\left[1-(1-\delta)^\theta\right]^{m-1}}\right)^{\frac{1}{\theta}}\right]$$

where $0 < \delta \leq 1$ and $\theta \geq 1$.
(Joe, 1997, pp. 153–154)

(7) The negative binomial Laplace transform

$$\phi(s) = \left[\frac{1-\theta}{\mathrm{e}^s - \theta}\right]^\delta$$

and its inverse

$$\phi^{-1}(t) = \ln\left[(1-\theta)t^{-\frac{1}{\delta}} + \theta\right]$$

generate the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \left[\frac{1-\theta}{\prod_{j=1}^m \left[(1-\theta)u_j^{-\frac{1}{\delta}} + \theta\right] - \theta}\right]^\delta$$

where $\delta > 0$ and $0 < \theta < 1$.
(Joe, 1997, p. 154)

C.3 Three dependence parameters

The power variance Laplace transform

$$\phi(s) = \mathrm{e}^{-\frac{\delta}{\alpha}[(\theta+s)^\alpha - \theta^\alpha]}$$

and its inverse

$$\phi^{-1}(t) = \left(-\frac{\alpha}{\delta}\ln[t] + \theta^\alpha\right)^{\frac{1}{\alpha}} - \theta$$

generate the multivariate Archimedean copula

$$C_{\mathrm{M}}(\mathbf{u}) = \mathrm{e}^{-\frac{\delta}{\alpha}\left(\left[\Sigma_{j=1}^{m}\left(\theta^\alpha - \frac{\alpha}{\delta}\ln[u_j]\right)^{\frac{1}{\alpha}} - (m-1)\theta\right]^\alpha - \theta^\alpha\right)}$$

where $\delta > 0$ and $\begin{cases} 0 < \alpha \leq 1 \\ 0 \leq \theta \end{cases}$ or $\begin{cases} \alpha \leq 0 \\ \theta > 0 \end{cases}$. This family includes as special case the copula obtained from an inverse Gaussian Laplace transform $(\alpha = \frac{1}{2})$, a gamma Laplace transform ($\alpha = 0$), and a positive stable Laplace transform ($\theta = 0$). (Hougaard, 1986a; Joe, 1997, p. 154 (for $\delta = \alpha$); Hougaard, 2000, pp. 241–243 and 504–506)

C.4 Bivariate extension to negative dependence

(1) The extended gamma Laplace transform is

$$\phi(s) = \begin{cases} e^{-s} & \delta = 0 \\ (1+s)^{-1/\delta} & \delta \neq 0 \text{ and } -1 < \delta \end{cases}$$

and its extended inverse is

$$\phi^{-1}(t) = \begin{cases} -\ln t & \delta = 0 \\ t^{-\delta} - 1 & \delta \neq 0 \text{ and } -1 < \delta \end{cases}$$

The bivariate Archimedean copula extended for negative dependencies is

$$C_{\mathrm{B}}(u_1, u_2) = \begin{cases} uv & \delta = 0 \\ [\max(u^{-\delta} + v^{-\delta} - 1, 0)]^{-1/\delta} & \delta \neq 0 \text{ and } -1 < \delta \end{cases}$$

Figure C.20 shows the bivariate dependence for $\delta = -0.3$ with uniform, Gaussian, and Laplace margins.
(Genest and MacKay, 1986a; Joe, 1997, pp. 157–158; Nelsen, 1999, pp. 94–95)

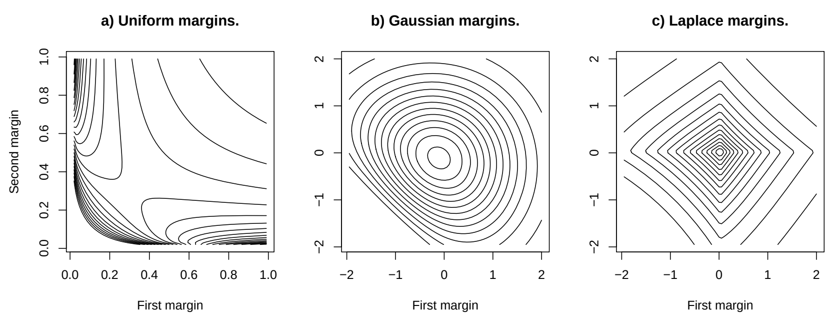

Fig. C.20. Density contours for a bivariate Archimedean copula generated from a gamma Laplace transform with a) uniform, b) Gaussian, and c) Laplace margins ($\delta = -0.3$).

(2) The extended logarithmic series Laplace transform is

$$\phi(s) = \begin{cases} e^{-s} & \delta = 0 \\ -\frac{1}{\delta}\ln\left[1 + e^{-s}\left(e^{-\delta} - 1\right)\right] & \delta \neq 0 \end{cases}$$

and its extended inverse is

$$\phi^{-1}(t) = \begin{cases} -\ln t & \delta = 0 \\ -\ln \frac{e^{-\delta t} - 1}{e^{-\delta} - 1} & \delta \neq 0 \end{cases}$$

The bivariate Archimedean copula extended for negative dependencies is

$$C_{\mathrm{B}}(u_1, u_2) = \begin{cases} uv & \delta = 0 \\ -\frac{1}{\delta}\ln\left[1 + \frac{(e^{-\delta u} - 1)(e^{-\delta v} - 1)}{e^{-\delta} - 1}\right] & \delta \neq 0 \end{cases}$$

and the multivariate Archimedean copula extended for negative dependencies is

$$\mathrm{C}(\mathbf{u};\delta) = \begin{cases} \prod_{j=1}^{m} u_j & \delta = 0 \\ -\frac{1}{\delta}\ln\left[1 + \frac{\prod_{j=1}^{m}(e^{-\delta u_j} - 1)}{(e^{-\delta} - 1)^{m-1}}\right] & \delta \neq 0 \end{cases}$$

Figure C.21 shows the bivariate dependence for $\delta = -5$ with uniform, logistic, and Burr ($\mu = 2$, $\sigma = 2$, And $\kappa = 2$) margins.
(Joe, 1997, pp. 158–159; Nelsen, 1999, pp. 94–95)

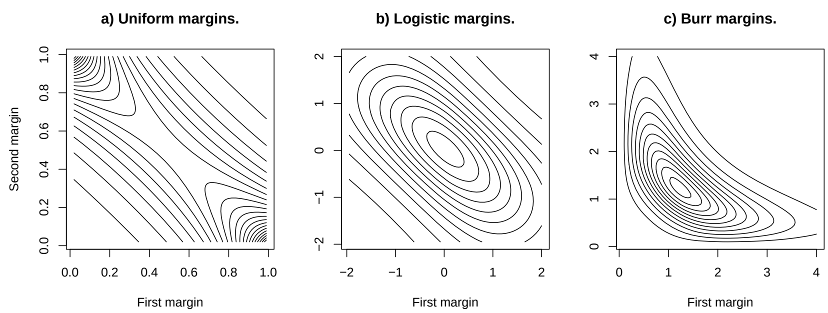

Fig. C.21. Density contours for a bivariate Archimedean copula generated from a logarithmic series Laplace transform with a) uniform, b) logistic, and c) Burr margins ($\delta = -5$).

Bibliography

1. Abramowitz, M. and Stegun, I.A. (1965) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover.
2. Agresti, A. (1989) A survey of models for repeated ordered categorical response data. *Statistics in Medicine* **8**, 1209–1224.
3. Agresti, A. (1990) *Categorical Data Analysis.* New-York: Wiley.
4. Agresti, A. (1999) Modelling ordered categorical data: recent advances and future challenges. *Statistics in Medicine* **18**, 2191–2207.
5. Agresti, A. and Lang, J.B. (1993) A proportional odds model with subject-specific effects for repeated ordered categorical responses. *Biometrika* **80**, 527–534.
6. Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. in Petrov, B.N. and Csàki, F., eds *Second International Symposium on Inference Theory*, Budapest: Akadémiai Kiadó, pp. 267–281.
7. Albert, P.S., Hunsberger, S.A., and Biro, F.M. (1997) Modeling repeated measures with monotonic ordinal responses and misclassification, with applications to studying maturation. *Journal of the American Statistical Association* **92**, 1304–1311.
8. Ali, H.M., Mikhail, N.N., and Haq, M.S. (1978) A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis* **8**, 405–412.
9. Barnard, G.A. (1954) Sampling inspection and statistical decisions. *Journal of the Royal Statistical Society B* **16**, 151–165.
10. Barnett, V.D. (1980) Some bivariate uniform distributions. *Communications in Statistics* **A9**, 453–461.
11. Bartholomew, D.J. (1980) Factor analysis for categorical data. *Journal of the Royal Statistical Society B* **42**, 293–321.
12. Bayes, T. (1958) An essay towards solving a problem in the doctrine of chances. *Biometrika* **45**, 296–315.
13. Clayton, D.G. (1978) A model for association in bivariate life-tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.
14. Conaway, M.R. (1989) Analysis of repeated categorical measurements with conditional likelihood methods. *Journal of the American Statistical Association* **84**, 53–62.
15. Conaway, M.R. (1990) A random effects model for binary data. *Biometrics* **46**, 317–328.
16. Cook, R.D. and Johnson, M.E. (1981) A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society B* **43**, 210–218.

17. Cox, D.R. (1972) Regression models and life tables. *Journal of the Royal Statistical Society B* **34**, 187–220.
18. Cox, D.R. and Oakes, D. (1984) *Analysis of Survival Data.* London: Chapman & Hall.
19. Crouchley, R. (1995) A random-effects model for ordered categorical data. *Journal of the American Statistical Association* **90**, 489–498.
20. Ezzet, F. and Whitehead, J. (1991) A random effects model for ordinal responses from a crossover trial. *Statistics in Medicine* **10**, 901–907.
21. Farewell, V.T. (1982) A note on regression analysis of ordinal data with variability of classification. *Biometrika* **69**, 533–538.
22. Frank, M.J. (1979) On the simultaneous associativity of F(x,y) and x+y-F(x,y). *Æquationes Math* **19**, 194–226.
23. Gange, S.J., Linton, K.L.P., Scott, A.J., DeMets, D.L., and Klein, R. (1995) A comparison of methods for correlated ordinal measures with ophthalmic applications. *Statistics in Medicine* **14**, 1961–1974.
24. Genest, C. and MacKay, R.J. (1986a) Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics* **14**, 145–159.
25. Genest, C. and MacKay, R.J. (1986b) The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician* **40**, 280–283.
26. Genest, C. and Ghoudi, K. (1994) Une famille de lois bidimensionnelles insolite. *C. R. Acad. Sci. Paris Sér. I Math.* **318**, 351–354.
27. Gibbons, R.D. and Hedeker, D.R. (2000) Applications of mixed-effects models in biostatistics. *Sankhyā: The Indian Journal of Statistics B* **62**, 70–103.
28. Gómez, E., Gómez-Villegas, M.A., and Marin, J.M. (1998) A multivariate generalization of the power exponential family of distributions. *Communications in Statistics* **A27**, 589–600.
29. Goodman, L.A. (1979) Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association* **74**, 537–552.
30. Goodman, L.A. (1981) Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association* **76**, 320–334.
31. Greenland, S. (1994) Alternative models for ordinal logistic regression. *Statistics in Medicine* **13**, 1665–1677.
32. Gumbel, E.J. (1960a) Bivariate exponential distributions. *Journal of the American Statistical Association* **55**, 698–707.
33. Gumbel, E.J. (1960b) Distributions des valeurs extrêmes en plusieurs dimensions. *Publication de l'Institut de Statistique de l'University de Paris* **9**, 171–173.
34. Hartzel, J., Liu, I-M., and Agresti, A. (2001) Describing heterogeneous effects in stratified ordinal contingency tables, with application to multi-center clinical trials. *Computational Statistics and Data Analysis* **35**, 429–449.
35. Harvey, A.C. (1989) *Forecasting, structural time series models and the Kalman filter.* Cambridge: Cambridge university press.
36. Hedeker, D. and Gibbons, R.D. (1994) A random-effects ordinal regression model for multilevel analysis. *Biometrics* **50**, 933–944.

37. Hougaard, P. (1984) Life Time Methods for Heterogeneous Populations: Distributions describing the heterogeneity. *Biometrika* **71**, 75–83.
38. Hougaard, P. (1986a) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 387–396.
39. Hougaard, P. (2000) *Analysis of Multivariate Survival Data.* New-York: Springer.
40. Ihaka, R. and Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, **5**, 299–314.
41. Jansen, J. (1990) On the statistical analysis of ordinal data when extravariation is present. *Applied Statistics* **39**, 75–84.
42. Joe, H. (1993) Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis* **46**, 262–282.
43. Joe, H. (1997) *Multivariate Models and Dependence Concepts.* London: Chapman & Hall.
44. Johnson, N.L. and Kotz, S. (1969) *Distributions in Statistics. Discrete Distributions.* New-York: Wiley
45. Johnson, N.L. and Kotz, S. (1970a) *Distributions in Statistics. Continuous Univariate Distributions - 1.* New-York: Wiley
46. Johnson, N.L. and Kotz, S. (1970b) *Distributions in Statistics. Continuous Univariate Distributions - 2.* New-York: Wiley
47. Johnson, N.L. and Kotz, S. (1972) *Distributions in Statistics. Continuous Multivariate Distributions.* New-York: Wiley
48. Jones, R.H. (1993) *Longitudinal Data Analysis with Serial Correlation: A State-space Approach.* London: Chapman & Hall.
49. Jørgensen, B. (1982) *Statistical Properties of the Generalized Inverse Gaussian Distribution.* New-York: Springer
50. Keen, A. and Engel, B. (1997) Analysis of a mixed model for ordinal data by iterative re-weighted REML. *Statistica Neerlandica* **51**, 129–144.
51. Kenward, M. and Jones, B. (1991) The analysis of categorical data from cross-over trials using a latent variable model. *Statistics in Medicine* **10**, 1607–1619.
52. Kenward, M.G., Lesaffre, E., and Molenberghs, G. (1994) An application of maximum likelihood and generalized estimating equations to the analysis of ordinal data from a longitudinal study with cases missing at random. *Biometrics* **50**, 945–953.
53. Kimeldorf, G. and Sampson, A.R. (1975) Uniform representations of bivariate distributions. *Communications in Statistics* **4**, 617–627.
54. Lapp, K., Molenberghs, G., and Lesaffre, E. (1998) Local and global cross ratios to model the association between ordinal variables. *Computational Statistics and Data Analysis* **28**, 387–411.
55. Lesaffre, E., Molenberghs, G., and Dewulf, L. (1996) Effect of dropouts in a longitudinal study: an application of a repeated ordinal model. *Statistics in Medicine* **15**, 1123–1141.
56. Lindsay, B.G., Clogg, C.C., and Grego, J., (1991) Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association* **86**, 96–107.

57. Lindsey, J.K. (1999) *Models for Repeated Measurements, 2nd edition.* Oxford: Oxford University Press.
58. Lindsey, J.K. and Jones, B. (1998) Choosing among generalized linear models applied to medical data. *Statistics in Medicine* **16**, 59–68.
59. Lindsey, J.K., Jones, B., and Ebbutt, A.F. (1997) Simple models for repeated ordinal responses with an application to a seasonal rhinitis clinical trial. *Statistics in Medicine* **16**, 2873–2882.
60. MacDonald, I.L. and Zucchini, W. (1997) *Hidden Markov and Other Models for Discrete-Valued Time Series.* London: Chapman & Hall.
61. McCullagh, P. (1980) Regression models for ordinal data. *Journal of the Royal Statistical Society B* **42**, 109–142.
62. Molenberghs, G. (1992) *A Full Maximum Likelihood Method for the Analysis of Multivariate Ordered Categorical Data.* Ph.D. thesis, Universiteit Antwerpen, Belgium.
63. Molenberghs, G. and Lesaffre, E. (1994) Marginal modelling of correlated ordinal data using a multivariate Plackett distribution. *Journal of the American Statistical Association* **89**, 633–644.
64. Molenberghs, G., Kenward, M.G., and Lesaffre, E. (1997) The analysis oflongitudinal ordinal data with nonrandom drop-out. *Biometrika* **84**, 33–44.
65. Nelder, J.A. and Wedderburn, R.W.M. (1972) Generalized linear models. *Journal of the Royal Statistical Society A* **135**, 370–383.
66. Nelsen, R.B. (1999) *An Introduction to Copulas.* New-York: Springer.
67. Pulkstenis, E.P., Ten Have, T.R., and Landis, R. (1998) Model for the analysis of binary longitudinal pain data subject to informative dropout through remedication. *Journal of the American Statistical Association* **93**, 438–450.
68. Ribaudo, H.J., Bacchi, M., Bernhard, J., and Thompson, S.G. (1999) A multilevel analysis of longitudinal ordinal data: evaluation of the level of physical performance of women receiving adjuvant therapy for breast cancer. *Journal of the Royal Statistical Society A* **162**, 349–360.
69. Schweizer, B. and Sklar, A. (1983) *Probabilistic Metric Spaces.* New-York: North Holland.
70. Salas, S.L. and Hille, E. (1990) *Calculus: One and several variables, 6th edition.* New-York: Wiley.
71. Seshadri, V. (1993) *The Inverse Gaussian Distribution. A Case Study in Exponential Families.* Oxford: Oxford University Press.
72. Song, P.X.K. (2000) Multivariate dispersion models generated from Gaussian copulas. *Scandinavian Journal of Statistics* **27**, 305–320.
73. Song, P.X.K. and Tan, M. (2000) Marginal models for longitudinal continuous proportional data. *Biometrics* **56**, 496–502.
74. Ten Have, T.R. (1996) A mixed effects model for multivariate ordinal response data including correlated discrete failure times with ordinal responses. *Biometrics* **52**, 473–491.
75. Wetherill, G.B. (1961) Bayesian sequential analysis. *Biometrika* **48**, 281–292.
76. Wilkinson, G.N. and Rogers, C.E. (1973) Symbolic description of factorial models for analysis of variance. *Applied Statistics* **22**, 392–399.